Peer Review File

Article information: http://dx.doi.org/10.21037/atm-20-4578

Reviewer A:

Major Comments and Reply:

Comment 1: How reliable can cancer-specific survival be determined in SEER?

Reply: The SEER registry is the largest population-based database of oncology patients in the United States, as per their report covering approximately 28% of patients diagnosed with cancer in the nation, including 13 regions. We defined the cancer-specific survival as the time from the date of diagnosis to the date of death caused by MCC. Many scholars have carried on the researches according to the CSS data in the SEER database and made some high visibility paper, such as‘Conditional survival of malignant thymoma using national population-based surveillance, epidemiology, and end results (SEER) registry (1973-2011)’ which was published on ‘Journal of Thoracic Oncology’, ‘Long-term survival among patients with Hodgkin's lymphoma who developed breast cancer: a population-based study’ published on ‘Journal of clinical oncology’ and ‘Survival outcome of local excision versus radical resection of colon or rectal carcinoma: a Surveillance, Epidemiology, and End Results (SEER) population-based study’ published on ‘Annals of Surgery’.

Comment 2: **Please provide more explanation as to why stage is an inadequate predictor of survival, and how a multivariate prediction model would change care of MCC patients. In other words, what do the authors envision to be the clinical application of this model?**

Reply: The introduction part has been revised according to the reviewer’s comment. The AJCC stage has played an important role in prognostic evaluations of tumors for a long time and can be used to divide patients into different risk categories. However, the AJCC stage only takes tumor size, lymph nodes and histological metastasis into account. There are some other prognostic factors affecting the survival rate of MCC patients , such as age at diagnosis, race, sex, marital status, surgical interventions.

Previous studies have found that these prognostic factors may affect the prognosis of MCC patients. Nomogram is a reliable and convenient tool for cancer prognosis prediction. As a result, we tried to use integrated nomograms that various prognosis prediction factors and developing personalized treatment options.

Change in the text: As shown in Line 69-78, page 4 in the manuscript. We elaborate the importance of constructing the nomograms.

Comment 3: **What is the purpose of 2 models (one for overall survival and one for MCC-specific survival)? Since the significant variables are different in each model, which do the authors recommend clinicians use?**

Reply: The CSS was calculated after classification according to MCC characteristics, especially referring to people who die of MCC. The overall- survival was defined as the time from surgery to death of any cause or, for living patients, to the date of last available information. The MCC was mainly found in elderly persons. With the increase of age, the physiological functions of the body decrease and more susceptible to disease which have great impact on survival. As a result, we use two kind of endpoints. The primary endpoint was 3- year CSS; the secondary endpoint was 3-year OS. The CSS is the most relevant indicator, but OS should not be ignored.

Change in the text: 1. Line 131-132, page 7. "The primary endpoint in our research was 3- year cancer-specific survival (CSS). The CSS was calculated after classification according to MCC characteristics, especially referring to people who die of MCC" was added.

2. Line 144, page 7. " The secondary endpoint in our research was 3-year Overall survival " was added.

3. Line 148-149, page 7. "The CSS is the most relevant indicator, but OS should not be ignored" was added.

Comment 4: **Since radiotherapy often differs in dose and treatment interval (aka palliative vs. curative intent), and since chemotherapy is a very heterogeneous category, has significant side effects and limited efficacy in MCC, treatment**

methods should be removed from the model entirely.

Reply: As suggested by reviewer, the radiotherapy and chemotherapy, which have significant impact on survival, need more details. This is definitely a correct suggestion, but unfortunately the SEER Program does not provide data about details of radiotherapy and chemotherapy. The chemotherapy and radiotherapy have effects on disease prognosis in patients.Therefore, we thought these details of the radiotherapy and chemotherapy belongs to the limitations of our research.

Change in the text: Line 294-298, page 13. "The variables obtained from SEER database were limited, and there were some important variable information that was not available, including the "No/Unknown" information, type, or extent of treatment and specific doses and drug regimens used. In addition, we did not account for length of time required for staging and treatment,which may lead to immortal time bias. " was added.

Comment 5: **Multiple predictor variables are strongly related to stage (e.g., size >2cm, LN involved, chemotherapy). The multivariate models may be affected by including these inter-related predictors in the same model.**

Reply: According to the reviewer's suggestion, we eliminate the inter-related prediction variables, including tumor size and lymph nodes. Then we present a reanalysis of revised data and obtain several new results in our multivariate models.

Change in the text: Multiple predictor variables which related to stage were eliminated from multivariate analysis. As shown in Table.2 and Table.3.

Comment 6: **What is the justification for combining "No" and "Unknown" for radiation and chemotherapy? These are very distinct categories.**

Reply: There are only 'no' and 'unknown' data in the SEER database,which we can't separate completely. As a result, we combine " No" and "Unknown" for radiation and chemotherapy. We have acknowledged this limitation in the Discussion.

Change in the text: Line 294-298, page 3-4. "There were some important variable information that was not available, including the "No/Unknown" information" was

added in the limitations.

Comment 7: The Cox model makes sense for overall survival, but for cancer-specific survival there is the competing risk of other causes of death. Hazard ratios for CSS can be estimated by censoring at the time of a competing event with the Cox model, but the probabilistic predictions are expected to be biased towards higher risk of MCC death. This can be seen in Figure 6A/B where the actual survival is higher than predicted survival on the left side of the curve. This does not appear to be as much of an issue in the validation set though (Fig 6C/D). A Fine & Gray competing risk regression model is generally preferred in this setting.

Reply: As the reviewer suggested, death from other causes is a competitive event of MCC - specific death. A Fine & Gray competing risk regression model describes the probability of occurring an event prior to a specific time based on cumulative incidence function (CIF). This model will not ignore the other competing risks associated with specific causes. We use the Fine and Gray competing risk regression model, which make our nomograms more reliably. Ultimately, we constructed a competing risk nomogram to cancer-specific survival.

Change in the text: Line 210-211, page 10. According to Fine and Gray analysis results, we drew the Fig.3.

Comment 8:**The type of ROC curves used and corresponding AUC statistics are for binary outcomes, not time-to-event outcomes with censoring. Instead, time-dependent curves should be used, for example Heagerty et al. Biometrics 2000;56(2):337-44 (https://pubmed.ncbi.nlm.nih.gov/10877287/).**

Reply: We reanalyze data according to the reviewer's suggestion. This reference have now been included into the cited literature. The C-index value of CSS and OS nomograms were calculated respectively which shown. Comparing with the the type of receiver operating characteristic curves we used before, time-dependent curves extend the concepts of sensitivity and specificity to time-dependent binary variables.

Time-dependent ROC curve considered not only the result-related but also the time-related factors. The using of time-dependent ROC curve for evaluating and comparing the prognostic capacity of diagnostic would greatly benefit validation of our models. It is useful in comparing our model with presently used AJCC staging system.

Change in the text: Line 150-152, page 7 and Table 4. "The nomogram accuracy was analyzed by Harrell C-index, together with the area under the time-dependent ROC curve" and results were added.

Comment 9:The decision curve analysis needs more elaboration. First, the reference provided is to another study that used the technique rather than to an adequate description and justification of the method. Second, the terms used like "standardized net benefit" are not defined, and thus are not currently interpretable. What does the magnitude of net benefit mean? How should these results be applied in practice? Are the DCA curves adding much to the other measures of performance provided? How does this translate to clinical practice?

Reply: The unsuitable references were changed. The decision curve analysis was first written in the paper entitled "Decision curve analysis: a novel method for evaluating prediction models" published in the 'Medical Decision Making'. We add new description content about the decision curve analysis in the methods. The AUC focuses on the predictive accuracy of a model. Decision-analytic methods incorporate consequences. The net benefit was calculated by subtracting the proportion of false positives from the proportion of true positives, and weighting by the relative harm of foregoing treatment compared with the negative consequences of an unnecessary treatment. The DCA is used to evaluate the performance of prediction models in clinical decision making, which could tell us which of several alternative models should be used. Clinicians have to decide whether interventions should be taken to the disease.

Change in the text: Line231-239, page 11. The elaborations of the DCA and the meaning of net benefit were added.

Comment 10:In addition, DCA appears to be have been proposed for binary outcomes. Are your calculations accounting for censoring of the outcomes?

Reply: DCA curves were commonly used in logistic models. In our research Cox regression analysis was used to assess the independent predictors for MCC which has censored values and follow-up time. The Vickers AJ extended the scope application of DCA to the analysis of survival data. We used this new method to solve the censoring of the outcomes. We cited the new literature "Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular marker. BMC Medical Informatics and Decision Making. 2008 Nov 26;8(1):53. s."

Change in the text: Line 231-239, page 11.

Comment 11:What was the median and range of follow-up (distinct from median survival)?

Reply: The median and range of follow-up has now been more carefully modified in the Table 1.

Change in the text: As shown in Table.1.

Comment 12:**"Survival time" and "median survival" are reported, but these appear to not have been calculated using the Kaplan-Meier method to account for censoring, so are mislabeled. These are actually descriptive statistics of the follow-up time. The median, inter-quartile range, and range of the follow-up time are more useful than the mean +/- SD because of the skewness of their distribution.**

Reply: As the Reviewer correctly said, the distribution of the follow-up time is skewness. We have made corrections in the our data analysis, adding the median, inter-quartile range of follow-up time into our statistical description.

Change in the text: As shown in Table.1.

Comment 13:**How many deaths were recorded? How many MCC-specific deaths and other cause deaths? Was cause of death unknown in any case? If so, how were those cases handled?**

Reply: There were 2010 patient deaths during the follow-up period: 645 cancer-specifific 10 deaths and 1365 competing mortalities. According to different AJCC stage, We use the Fine and Gray competing risk regression models to handle these death,including non-cancer-specific death and cancer-specific death .

Change in the text: Line 176-182, page 8-9. The supplementary documents gave details of Fine and Gray competing risk regression models.

Comment 14:**It would be helpful to include figures showing survival in each stage group, using the Kaplan-Meier estimator for overall survival and cumulative incidence curves for MCC-specific death, accounting for the competing risk of other causes of death. Some of the other figures can be moved to supplemental material or removed, especially performance on the training set which is less relevant than validation set performance.**

Reply: Based on different AJCC stage, the competing risks of cancer - specific deaths and other mortalities were analyzed with the Fine and Gray proportional hazards model. According AJCC stage, We plotted figures to show the cumulative incidence of cancer-specific death and other-diseases death. In our analysis, the cumulative incidence of non-cancer-specific death was higher than that of cancer-specific death in patients other than AJCC IV stage. As suggested these figures have been moved to supplementary documents or removed.

Change in the text: The supplementary documents gave details of Fine and Gray competing risk regression models.

Minor Comments and Reply:

Comment 1: In the abstract, avoid the term “internal validation.” That refers to the

use of a specific set of techniques like cross-validation, rather than simply estimating performance using the same data used to train the model [Han K, Korean J Radiol, 2016;17(3):339-350. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4842854/]

Reply: We have modified according to the reviewer's suggestion.

Comment 2: typo'aetilogy', correct to'etiology'(line 56, page 3)

Reply: It has been modified in the introduction section.

Comment 3:Why were patients with unknown primary tumors excluded? "excluded...primary site not on skin" (Line 111, page 6)

Reply: Merkel cell carcinoma is a rare and aggressive skin cancer and classified as neuroendocrine tumours of the skin. The biological effects of the mucosa may be inconsistent with those of the skin. Therefore, We analyze Merkel cell carcinoma that primary in the skin.

Comment 4: The authors did not account for length of time required for staging and treatment (often this takes months), leading to possible immortal time bias. If this is not possible with SEER dataneeds to be mentioned in limitations.

Reply: These would be helpful additions. We have added this point in the limitations.

Change in the text: Line297-298, page 13-14.

Comment 5: **Chemotherapy is not recommended as first line therapy for MCC for AJCC stages I-III. Therefore, the model would be more clinically useful if chemotherapy was removed.**

Reply: The reviewer's suggestion for the Chemotherapy is correct and we are grateful for this suggestion. Removing the chemotherapy from our analysis may affect the accuracy and stability of our results. There are also some new chemotherapy regimens were administered, which may have impact on our analysis. We have added this point in the limitations.

Change in the text: Line294-297, page 13.

Reviewer B:

Comment:This manuscript entitled "Nomogram Prediction for the Overall Survival and Cancer-specific Survival of Patients Diagnosed with Merkel Cell Carcinoma" by Xufeng YIN reported the prognostic factors for patients with Merkel Cell Carcinoma (MCC) and performed nomogram prediction using SEER database. Generally, the idea is good and statistical analysis was well-performed. However, the authors may lack clinical insight which limit the novelty and reliability of this nomogram. In addition, the manuscript was not well-organized and well-written by English and many grammar errors could be found. Furthermore, the authors misunderstood some references and cited incorrect references to support they findings.

Reply: We tried our best to improve the manuscript and made some changes in the manuscript. And here we marked in red in revised paper. In addition, the wordings of the main text, figures and tables have been checked by a native English-speaking expert who is majoring in dermatological field. We appreciate for Reviewer's warm work earnestly, and hope that the correction will meet with approval. According to the reviewers' suggestions, we use a Fine & Gray competing risk regression model describes the probability of occurring an event prior to a specific time based on cumulative incidence function. We have revised the literature to make sure it fully support our manuscript.We hope these modifications helped to elevate the quality of the article.Thank you very much for your comments and suggestions.