Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-24-1837

Reviewer A

Comment 1: The study would benefit from an external validation cohort to confirm the generalizability of the proposed nomogram. The authors could discuss the feasibility of such validation in the future. If possible, please include validation with external or multi-center datasets to enhance the model's reliability

Reply1: Thank you very much for the positive feedback and constructive points. Thymic epithelial tumors are rare tumors, and currently our center does not have sufficient cases to conduct external validation, which is a limitation of our study. We hope to use multi-center datasets for external validation and further enhance the model's reliability in the future (see Page 9, line 255-258).

Changes in the text: "a large-scale external validation cohort with multi-center datasets is needed to further confirm prognostic predictivity of this nomogram model." was written. (see Page 9, line 255-258).

Comment 2: The C-index of 0.68 indicates only moderate predictive performance. The manuscript would benefit from comparing the nomogram's performance with other predictive methods, such as machine learning models, to contextualize its accuracy and reliability. Could the authors compare the nomogram's performance with other advanced predictive models and explore ways to improve its accuracy? Additional metrics such as area under the ROC curve (AUC) could provide a clearer understanding of its discriminative power.

Reply 2: Thank you for your insightful suggestion. We think this is an excellent suggestion. However, no other prediction models for metastatic thymic epithelial tumors are reported and we are currently unable to compare with other models. We plan to combine genetic testing results to compare models in the future. And we have modified our text as advised (see Page 6, line 151-153) by adding area under the ROC curve (AUC), which provide a clearer understanding of nomogram's discriminative power.

Changes in the text: "We also utilized the ROC curve to assess the performance of the nomogram. The area under the ROC curve(AUC)values at 2 and 3 years were 0.74 and 0.73, respectively, indicating that the nomogram exhibited good predictive accuracy.(Figure 2A, 2B)" was added. (see Page 6, line 151-153)

Comment 3: The clinical significance and role of key variables in the nomogram are insufficiently explained. For example, how each variable contributes to the prediction is unclear. How about providing a detailed discussion of the clinical implications of the significant variables and their relative weights in the model?

Reply 3: Thank you for your professional suggestion. According to your suggestion, we explained how each variable contributes to the prediction.

The WHO histologic classification: (see Page 7, line 181-186)

Masaoka-Koga stage: (see Page 7, line 203-207)

KPS score: (see Page 8, line 218-221)

baseline serum albumin level: (see Page 8, line 226-232)

We also added a discussion of the clinical implications of the significant variables and their relative weights in the model (see Page 8, line 235-244).

Changes in the text: "There are no truly benign epithelial tumors of the thymus, but different subtypes exhibit varying biological behaviors and levels of clinical aggressiveness. The metastatic potential varies among thymic tumors, escalating from type A and AB thymomas, through B1, B2, and B3 thymomas, to the most aggressive form, thymic carcinomas. This variation in metastatic potential leads to different treatment strategies and follow-up schedules after diagnosis" was added (see Page 7, line 181-186).

"For resectable thymic TETs, complete surgical resection is recommended. Radiotherapy or chemotherapy can be employed as postoperative adjuvant treatment. In cases of advanced or unresectable TETs, the primary treatments are radiotherapy and chemotherapy [21]. The different treatments lead to different prognoses" was added (see Page 7, line 203-207).

"WHO histologic classification, Masaoka-Koga stage, KPS score, and baseline serum albumin level were all independent prognostic factors of OS in patients with metastatic TETs in our study, but their relative weights in the model were totally different. In terms of weight, the histological type holded the greatest significance, followed by albumin levels and KPS scores, with the stage carrying the least weight. The possible reasons are as follows. Histological classification is associated with the aggressiveness of the tumor, and tumor aggressiveness is an important factor in determining prognosis of tumor patients. Staging is also a prognostic factor for tumors, but in this study, all included patients were in stage IV, and most of the treatments were palliative, resulting in less significant differences in prognosis, hence the lowest weight." was added (see Page 8, line 235-244).

Comment 4: The manuscript highlights the clinical utility of the nomogram but would benefit from specific examples or scenarios where this tool could directly influence treatment decisions in high-risk patients. Providing specific scenarios where the nomogram could guide clinical decisions will highlight its practical application, particularly in high-risk patients.

Reply 4: Thank you for your suggestion. Below, I will provide a specific scenario. For example, if a patient has thymic squamous cell carcinoma, is at stage IVA, has a KPS score of 70, a baseline serum albumin level of 35 g/L, then the patient's scores for each category would be 100/0/50/50, with a total score of 200. The estimated 2-year survival rate is 48%, and the 3-year survival rate is around 38%. A score more than 112 indicates the patient being in a high-risk group. According to domestic and international guidelines, the standard first-line systemic treatment is chemotherapy. Based on reports from the 2024 ESMO ASIA conference, the combination of chemotherapy and immunotherapy for the first-line treatment of thymic tumors can improve the objective response rate. Therefore, for this patient, if conditions permit, it is advisable to actively consider a combination of chemotherapy and immunotherapy. The patient should receive full dose chemotherapy and the chemotherapy interval should not be extended. It is important to promptly re-evaluate the effectiveness after two cycles of treatment.

Changes in the text: There is no need to change the text.

Comment 5: As the study was conducted in a single institution, the patient population might be skewed, potentially limiting the applicability of the findings. Please note a detailed description of the study population and discuss potential selection biases.

Reply 5: Thank you for your suggestion. A detailed description of the study population has been added (see Page 4, line 90-92). We have added a discussion on potential selection bias (see Page 9, line 253-255) .

Changes in the text: “For the purpose of this study, patients were included if they had (1) histologically confirmed thymic epithelial tumors; (2) distant metastases; and (3) palliative anti-tumor treatment. Patients with previous malignancy were excluded. ” was added. (see Page 4, line 90-92)

“Since this study is a retrospective study conducted at a single center, selection bias is inevitable. Future studies can be conducted at multiple centers to reduce the occurrence of bias.” was added (see Page 9, line 253-255)

Minor Comment 1: Consider reformatting Table 2 for better readability, particularly by highlighting statistically significant variables.

Reply: Thank you for your suggestion. We have reorganized Table 2 and highlighted the statistically significant variables.

Changes in the text:

Table 2 (part of table 2)

KPS score			**0.010**
90-100	56.57（38.54-74.59）	1.00	
70-80	31.87（5.71-58.02）	1.73（1.05-2.84）	
WHO histologic classification			**0.001**
A-AB	102.37（67.05-137.68）	1.00	
B1-B3	64.20（50.10-78.30）	1.74（0.82-3.72）	0.151
C	34.70（24.26-45.14）	3.15（1.48-6.73）	0.003
Masaoka-Koga stage			**0.013**
IVa	68.33（49.8-86.80）	1.00	
IVb	42.93（24.59-61.28）	1.62（1.11-2.37）	

Albumin			**0.001**
<40 g/L	32.10（20.43-43.77）	1.00	
≥40 g/L	65.47（47.27-83.67）	0.52（0.33-0.81）	

Minor Comment 2: Ensure that figures, especially calibration plots and decision curve analyses, have descriptive legends to aid readers in interpretation.

Reply: Thank you for your suggestion. We added detailed descriptive legends in Figure 2.

Changes in the text:

Descriptive legends in **Figure 2**

Figure 2 Receiver operating characteristic curve analysis for the sensitivity and specificity of the nomogram-predicted 2-year (A) and 3-year (B) overall survival. Calibration curves of the nomogram-predicted 2-year (C) and 3-year (D) overall survival. Nomogram-predicted OS is plotted on the x axis; actual OS is plotted on the y axis. A plot along the 45-degree line would indicate a perfect calibration model in which the predicted probabilities are identical to the actual outcomes. Decision curve analysis of the nomogram-predicted 2-year (E) and 3-year (F) overall survival. The y-axis represents the net beneft, and the x-axis represents the corresponding risk threshold. The green dashed line represents all patients who died during the follow-up period. The blue dashed line represents no patient deaths during the follow-up period. The red solid line represents the net beneft of the nomogram at diferent risk thresholds.

Reviewer B

Major issues are:

- There is not a validation cohort

- Patients have been treated as far back as 2020, and advances in systemic therapy (particularly immunotherapy and/or antivascular drugs for thymic carcinoma) make the results hardly applicable at present.

- The risk factors found (WHO histologic classification, Stage and Performance status) are already well known risk factors and add nothing to what is already considered in assessing the prognosis of this population.

Reply: Thank you for your comments.

-The overall incidence of thymic epithelial tumors is low, roughly one percent of the incidence of lung cancer, with a very small number of new cases each year. And our study is a single-center study, we currently do not have enough cases to support external validation, which is a limitation of our study, and it is also mentioned in the text (see Page 9, line 253-258).

-The NCCN guidelines have not seen many updates in recent years regarding the medical treatment of metastatic thymic epithelial tumors, especially in first-line treatment. Currently, anti-angiogenic drugs such as sunitinib and immune checkpoint inhibitors such as pembrolizumab are only recommended by the guidelines for subsequent-line treatment of thymic carcinoma, not for first-line treatment. Moreover, these drugs currently only have phase II clinical studies and lack phase III randomized controlled trials (RCTs). Currently, platinum-based chemotherapy remains a classic treatment regimen. With the emergence of new drugs, an increasing number of clinicians are likely to attempt to combine anti-angiogenic therapy or immunotherapy with first-line chemotherapy. In the future, we plan to study the impact of these drugs on the prognosis of thymic tumors, and then assess whether our prognostic model needs to be revised.

-We also fully agree that WHO histologic classification, stage and performance status are known prognostic factors for tumors. Our goal is to use common, cost-effective, and easily accessible prognostic factors to build a model to help assess patient prognosis. In the future, we will try to identify more new prognostic factors.