Peer Review File

Article information: https://dx.doi.org/10.21037/tcr-23-1451

Reviewer A

Comment to the author：Very well written article. The topic is excellent. Graph and tables were also well done. Highly recommend this article for publication.
Reply: We feel great thanks for your professional review work on our paper. Thank you for acknowledging our work.

Reviewer B

Comment to the author：
This is a cohort study including patients from the SEER database. The objective was to predict the risk of LNM in patients with pT1 or pT2 colon cancer and based on these risk factors to develop a prediction nomogram to help guide treatment selection.
The AUC of the model was 0.65 and the calibration was good.

Advantages
This is a well written manuscript that follows the tripod statement. The number of patients in the cohort is large, and the reported outcome in regards of the overall risk of LNM and AUC of the prediction tool is similar to what other studies have reported. The methodology for the statistical analysis seems correct.
Reply: Thank you very much for your time involved in reviewing the manuscript. We appreciate your clear and detailed feedback and hope that the explanation has fully addressed all of your concerns. Here is a point-by-point response to your comments and concerns.

Comment 1: The major drawback in this study is the lack of clinical utility. The study concerns patients with pathological stage tumors, but the decision for treatment is based on the clinical stage. The discrepancy between clinical and pathological staging is well known and reported in multiple studies on multiple populations, and makes the model in this study clinically unusable. The authors have not mentioned this major drawback in the discussion section.
Reply 1: Thank you for pointing this out. We agree with this comment. And we have modified our manuscript as advised (see Page 12-13, line 236-240).
Changes in the text: Unfortunately, because the discrepancy between clinical and pathological staging, it may cause a problem that the model in our study clinically unusable. To deal with this, the T staging and distant metastasis can be evaluated by routine enhanced magnetic resonance imaging or CT before surgery to maximize the consistency of clinical and pathological staging (24).

Comment 2: Some of the patients will have cT3-4/cN1-2 but in fact pT1-2/pN0 at histopathological examination. These patients have not been removed from the dataset since it does not contain information on clinical staging and represents a bias.
Reply 2: Agree. The corresponding discussion about this problem has been incorporated into the manuscript (see Page 13, line 240-242).
Changes in the text: Furthermore, the SEER database lacks clinical staging information, which may introduce bias if patients with discrepancies between clinical and pathological staging are not excluded(25).

Comment 3: The AUC is 0.65 (the training set is always a little higher). This is not a very high AUC and the authors should be very carefull not to put to much value in to this. In the present-day pathology, other pathological analysis performed on the tumors (in case of local resection) would often include dMMR, budding etc. This would probably give a better risk estimation than the data provided in the SEER database.
Reply 3: We agree with this and have incorporated your suggestion into the manuscript (see Page 10-11, line 192-194 and Page 12, line 231-233).
Changes in the text:
The AUC in this paper, although not exceptionally high, aligns consistently with several prior studies. Moreover, the DCA demonstrated good clinical utility in the proper range.

Secondly, we failed to use genetic markers and pathological data which including dMMR, budding and other parameters for model construction, because the SEER database lacks relevant data.

Reviewer C

Comment 1: Thank you for hard work in analyzing so much data. However, Distant metastasis should be excluded in this study as it predicts T1 and T2 lymph node metastasis.
Reply 1: Thank you for this suggestion. It would have been interesting to explore this aspect. However, in the case of our study, it seems slightly out of scope because the presence of distant metastasis was regarded as an independent risk factor by means of univariate and multivariate logistic analyses. Besides, the exclusion of patients staging pT1-2N0-2M1, who constituted a substantial proportion of our study population, may introduce selection bias into our findings. Thank you once again for your diligent efforts in reviewing our article.

Comment 2: More than 0.7 was considered to be useful generally, however AUC in this study was about 0.65, so the author should discuss about this point.
Reply 2: Thank you for pointing this out. We agree with this comment. Therefore, we have discussed this point in the revised manuscript (see Page 10-11, line 192-194).
Changes in the text: The AUC in this paper, although not exceptionally high, aligns consistently with several prior studies. Moreover, the DCA demonstrated good clinical utility in the proper range.