Peer Review File

Article information: http://dx.doi.org/10.21037/tau-21-49

Reviewer A: This paper describes the development of a nomogram (multivariable logistic regression) based on the variables PI-RADS and PSAD, that specifically seeks to identify significant cancers in the Transition Zone (TZ) of the prostate, using biparametric (no contrast) MRI.
The sample size is 383 and with a sufficient number of 74 events (clinically significant cancer in TZ) for two predictors in the model. The methods are overall sound and the authors report discrimination (AUC) and calibration plots. As is typical with the development of prediction models, the authors appropriately split the dataset in training and validation cohorts. The methods (imaging, pathology etc.) seem rigorous. The manuscript is well-written and strengthened by the reporting of the inter-rater reliability between 3 radiologists (kappa).
The study is limited by: being retrospective rather than prospective (risk of selection bias), lacking external validation (performance characteristics limited to the patient population under study), possible model overfit (high AUCs) and lacking decision analytical results.

MAJOR
Comment 1: The main limitation of the paper is the lack of external validation in a separate cohort. What are the plans for this? Please include this in the Discussion and consider revising the Conclusion: "Our study constructed a well-validated nomogram…"
Reply 1: Thank you very much for your comments and suggestions. Our study constructed a novel internally validated nomogram to predict the probability of cs-PCa in the TZ based on the PI-RADS v2.1 score and PSAD. Adding PSAD to the PI-RADS v2.1 could improve diagnostic performance, thereby avoiding unnecessary biopsies. However, the lack of external validation was indeed one of the main limitations of our study, which has been mentioned in the limitation section. We are collecting more independent samples for external validation, and the prediction model would be further verified in the future study. Furthermore, we have modified the Conclusion.
Changes in the text: see Page 18, line 2.

Comment 2: Please comment on the possibility of model overfit. The AUCs are very high (>0.90) and the model is only validated internally (on the same dataset). Please see rule "5.8. Correction for overfit is strongly recommended for internal validation" in the Assel et al guidelines for reporting statistics in for clinical research in urology (available for example at https://bjui-journals.onlinelibrary.wiley.com/doi/full/10.1111/bju.14640).
Reply 2: Thank you very much for your comments. This article was retrospective and single-center study, so patient selection bias may exist, which probably has contributed to model overfit. Multicenter studies would be carried out to further

validate the present results. However, in our study, the prediction model has been internally verified, showing a good agreement between the actual outcomes and predicted outcomes, and we are collecting more independent samples for further external validation.

Comment 3: From the description in the statistical analysis section of "Multivariate logistic regression analysis of cs-PCa in the TZ", did the authors do stepwise selection by only including the significant variables on univariate analysis in the multivariable model? Please see guideline rule 5.2. Avoid stepwise selection. Could this possibly have contributed to overfit also? Which variables did the authors hypothesize would be associated with csPCa based on clinical knowledge and for consideration of inclusion in the model?

Reply 3: Thank you very much for your comments and suggestions. We have revised the relevant parts in the manuscript. The determination of covariates is indeed crucial when it comes to developing a prediction model. Introducing confounding variates may reduce the accuracy of the model. In the revised manuscript, before multivariate logistic regression analysis, five easily-obtained clinical indicators were selected as candidate variables in this study. To avoid incorporating a confounding factor and better handle strongly-correlated variables, the multivariable model eliminated t-PSA and PV with smaller AUC ($AUC_{t\text{-}PSA}$=0.756, AUC_{PV}=0.299) compared with PSAD (AUC_{PSAD}=0.842), which has also reported in the previous study (Zhang *et al*., Cancer Manag Res, 2020. DOI:10.2147/CMAR.S250633). The final variables including age, f/t-PSA, PSAD and PI-RADS v2.1 score were incorporated into the multivariate logistic regression analysis, and then the independent predictors for cs-PCa in the TZ were determined.

Changes in the text: see Page 12, line 14-22.

Comment 4: Is the nomogram included in Figure 2 ready for clinical use given the limitations mentioned? If it is not ready for "prime time", should it be omitted from the paper?

Reply 4: Thank you very much for your comments and suggestions. The nomogram was the perfect example of integrating multiple related parameters to predict a specific end point by means of geometry graphs to visualize the results of the prediction model. With the nomogram, we could intuitively obtain a patient's corresponding risk value for TZ cs-PCa on the prediction line at the bottom of the nomogram, which can then guide urologists' decision on whether to conduct a biopsy. The newly constructed nomogram has been internally verified, external validation would be carried out in the further study. Hence, although there are some limitations, the nomogram still has some advantages.

Comment 5: Did the authors consider including a Decision Curve Analysis to determine the net benefit of the model vs PI-RADS or PSAD alone across a range of threshold probabilities? More details can be found here:

https://www.mskcc.org/departments/epidemiology-biostatistics/biostatistics/decision-curve-analysis

Reply 5: Thank you very much for your comments and suggestions. We have conducted a Decision Curve Analysis and added the decision curve in this revised manuscript. The decision curve analysis was presented in Fig.3E, 3F. The net benefit of the prediction model was higher than the individual predictors when the threshold probability exceeded 0.25 in the training cohort and 0.50 in the validation cohort, which indicates that the prediction model has more advantages in guiding physician decision-making.

Changes in the text: see Page 11, line 10-14; Page 13, line 16-19; Page 17, line 10-14; figure 3.

Comment 6: The authors mention that the model can reduce unnecessary biopsies but do not quantify this. Did the authors consider reporting the clinical consequences of using the nomogram, for example #csPCa caught and missed and #biopsies avoided per 1000 men? See Assel et al guidelines rule 5.11. Report the clinical consequences of using a test or a model. The authors could also consider reporting the net reduction in biopsies by comparing the nomogram to the individual predictors or a base model based on clinical variables (e.g. PCPT risk calculator or similar). This can also be plotted using decision analysis methods described under Decision curve analysis.

Reply 6: Thanks very much for your comment and suggestions. In this revised manuscript, we have modified relevant parts. The percentage of patients avoiding unnecessary biopsy have been added in table 4. The decision curve analysis was also presented in Fig.3E,3F. The net benefit of using the prediction model was higher than the individual predictors in most ranges of the threshold probabilities.

Changes in the text: see table 4; Page 11, line 10-14; Page 13, line 16-19; Page 17, line 10-14; figure 3.

Comment 7: Please justify the rationale for the chosen reporting guideline. The authors write: “We present the following article/case in accordance with the CONSORT reporting checklist.” However the CONSORT checklist pertains to randomized controlled trials and this is not an RCT. Further down in the manuscript it says: “The authors have completed the STROBE reporting checklist” which applies to observational studies and would be more appropriate. At the end of the manuscript, a TRIPOD Checklist: Prediction Model Development is included which is appropriate since this is reporting on a prediction model, however, it is not filled out. Reporting guidelines can be reviewed and downloaded from the Equator website: https://www.equator-network.org/

Reply 7: Thank you very much for your comments. We are so sorry for making a mistake in the reporting checklist, we have made changes in the manuscript. We have completed the TRIPOD reporting checklist.

Changes in the text: see Page 7 line 18; Page 18 line 20.

Comment 8: Why is the focus and outcome of the model TZ disease? What about csPCa in the PZ?
Reply 8: Thank you very much for your comments. Identification of transition zone lesions remains a challenge, as PCa in the TZ is susceptible to being affected by benign prostatic hyperplasia and other factors in MRI. The number of false-positive and false-negative cases increases, resulting in high tumor misdiagnosis rates in the TZ. In addition, PI-RADS v2.1 mainly modify the diagnostic criteria in the transition zone compared with PI-RADS v2. So we selected TZ PCa as the study objectives.

MINOR
Comment 1: Table 1. Are all the csPCa TZ cancers? If so, please specify to help the reader.
Reply 1: Thank you very much for your comments. We are so sorry for the misunderstanding due to our unclear description. Our study is focus on TZ lesions, we have revised the relevant parts in Table 1.
Changes in the text: see table 1.

Comment 2: Table 1. For continuous variables, are numbers reported median and IQR? If so please specify. Please also consider the level of precision in all tables. See Assel et al Guidelines rules 4.3. For descriptive statistics, median and quartiles are preferred over means and standard deviations (or standard errors); range should be avoided and 4.1. Use appropriate levels of precision
Reply 2: Thank you very much for your comments and suggestions. For the numbers of continuous variables, we have reported median and quartile. We are sorry for not specifying in the table. We have made specifications according to your comment in Table 1.
Changes in the text: see table 1.

Comment 3: Consider omitting the ROC figures. They rarely, if ever, tell us more than the AUC.
Reply 3: Thanks very much for your comment and suggestions. In this revised manuscript, we have revised the relevant parts of the paper according to your suggestions. The ROC figure has been omitted.

Comment 4: Spelling should be “multivariable” not “multivariate”.
Reply 4: Thanks very much for your comment and suggestions. In this revised manuscript, we have changed “multivariate” to “multivariable” according to your suggestions.

Reviewer B: The running title: A Nomogram Predicting CS-PCa in the Transition Zone, met the objectives and conclusions of the publication. The checklist of the reporting guidelines was correctly filled in. The theme of the article is relevant, with PI-RADS being studied in Transitional Zone. Bp-MRI was used in all patients,

which is also relevant because it can decrease time and costs in diagnosis. The study involved 511 patients, 74 (14,48%) of whom had cs-PCa, which is compatible with a single institution retrospective study. Statistical analysis showed PI-RADS v2.1 and PSAD as significant independent predictors for cs-PCa in the TZ. With these two independent predictors a nomogram was successfully constructed.

Comment 1: Figures and tables were relevant and well-constructed. One comment on table 3: the PI-RADS v2.1 P value = 0.000 could be replaced by <0.001 which is more frequent in literature.
Reply 1: Thanks very much for your comment and suggestions. We have made modifications in the manuscript as advised.
Changes in the text: see table 3.

Reviewer C
Comment 1: Please report the result according to the common standards of statistical analysis and reporting (especially tables) http://dx.doi.org/10.1016/j.eururo.2014.06.024
- P-Value can not be 0.000 as stated in table 3. Report p-values to a single significant figure unless the p-value is close to 0.05, in which case, it is reasonable to report two significant figures. There is no need to report p-value of 0.703 using three decimal places, p=0,7 is enough
- Hazard and odds ratios are normally reported to two decimal places, although this can be avoided for high odds ratios (eg, 18.2 rather than 18.17).
Reply 1: Thanks very much for your comment and suggestions. We have made changes in the manuscript as advised.
Changes in the text: see table 3.

Comment 2: Please redesign table 1 to make it more intuitive, present clear numbers and detection rates.
As a hint for PIRADS reporting, please use the form Tran et al https://doi.org/10.1148/radiol.2018170425 used in Table 3 of their publication "PI-RADS Version 2 Scoring of Transition Zone Lesions and Proportion of Cancer Detection of their publication." (E.g. crosstab PIRADS vs Gleason Score detected)
Reply 2: Thanks very much for your comment and suggestions. We have made modifications in the manuscript as advised.
Changes in the text: see table 1.

Comment 3: "Age, t-PSA, percent free PSA, PSAD, PV and PI-RADS v2.1 score as univariate indicators were incorporated in multivariate logistic regression analysis, and then the independent predictors for cs-PCa in the TZ were determined. "PSAD is a combined variable that is calculated from PV and PSA, so you can not include the combined variable together with its components into the same model. It is statistically wrong!

Reply 3: Thank you very much for your comments. We have revised the relevant parts in the manuscript. Introducing confounding variates may reduce the accuracy of the prediction model. Strong correlations exist between several variables (f/t-PSA was derived from t-PSA. PSAD was calculated by dividing t-PSA by PV). To avoid incorporating a confounding factor and better handle strongly-correlated variables, the multivariable model eliminated t-PSA and PV with smaller AUC ($AUC_{t\text{-}PSA}$=0.756, AUC_{PV}=0.299) compared with PSAD (AUC_{PSAD}=0.842), which has also reported in the previous study (Zhang et al., Cancer Manag Res, 2020. DOI:10.2147/CMAR.S250633). The final variables including age, f/t-PSA, PSAD and PI-RADS v2.1 score were incorporated into the multivariate logistic regression analysis, and then the independent predictors for cs-PCa in the TZ were determined.
Changes in the text: see Page 12, line 14-22.

Comment 4: Wibulpolpraser et al https://www.ajronline.org/doi/10.2214/AJR.19.21608 reported on the detection of prostate cancer lesions in the midgland, base, and apex. Does your model show constant accuracy disregarding the location of the lesion?
Reply 4: Thank you very much for your reasonable comments and suggestion. Our study was not divided the lesions into the midgland, base, and apex because of limited samples. Further researches would be carried out based on expanded population samples.

Comment 5: Jeong et al https://doi.org/10.1016/j.juro.2017.01.073 reported on the unfortunate effect of obesity on the number of transitional zone tumors. Does BMI achieve any level of significance if included in your model?
Reply 5: Thank you very much for your comments and suggestion. Our study data has not collected BMI and patients' BMI were rarely recorded in our institution. As for your valuable suggestions, we will consider them in future studies.

Comment 6: In a rather historical series Djavan et al reported on the influence of transition and total prostate volumes on prostate cancer yield https://pubmed.ncbi.nlm.nih.gov/10895015/ Please report on the mean TZ volumes of your cohort and whether higher TZ volume influences the model accuracy.
Reply 6: Thank you very much for your comments and suggestion. Our study data has not collected TZ volumes. Five easily-obtained clinical indicators have been selected as candidate variables in this study. As for your valuable suggestions, we will consider them in future studies.

Comment 7: I would recommend to review the manuscript to correct the writing and grammar. Eg Methods page 7, lines 9-10: "In addition, there were 63 patients who met the exclusion criteria, the remaining 511 patients were finally enrolled who were randomly split into 383 patients in the training cohort and 128 patients

in the validation cohort."
This sentence makes no sense, please rephrase.
Reply 7: Thank you very much for your comments and suggestion. We have rephrased this sentence in the relevant parts of the paper.
Changes in the text: see Page 8, line 10-12.