Peer Review File

Article Information: https://dx.doi.org/10.21037/tau-22-538

Reviewer A

In this study, the authors analyzed IHC of 14 markers on the tissue microarrays from a total of 236 patients diagnosed with urothelial carcinoma of the bladder in their own institution. They first applied CK5/6, CK14, GATA3 and CK20 to classify the bladder cancer specimens into luminal vs basal, and found that patients with basal tumor were related with poorer prognosis compared to those with luminal tumor. They further revealed luminal MIBC patients could be further categorized into subgroups by YAP1 and CCNB1 stains as independent risk factor. Based on these findings, the authors concluded that molecular subtyping into luminal/basal subtype and risk stratification method using a 2-marker method by immunohistochemistry can be an economical, clinically practical method to predict patient prognosis and could help to develop treatment strategy and follow-up schedule in clinical practice. The data is well presented, and the manuscript was well written.

Manor comments:

Comments 1. For the luminal/basal subtyping, the cut-offs were defined by pathologist, not generated from the training cohort. There is no sense to separate training cohort and validation cohort. The cut-offs of other markers (CCNB1 and YAP1) were generated from the training cohort, so the validation cohort made sense.

Reply 1. The classification of luminal/basal BCa by IHC were investigated by other researchers (PMID: 27612592, 31494492, 22308455, 31240474). Although the markers were similar, the methods vary widely between researches and no method are able to serve as reference. In this research, urologists and pathologists collaborate together to explore the molecular classification method based on IHC. The IHC slides of different markers were viewed by pathologist and IHC score of each marker were given. While urologists combined the IHC score and clinical characteristics to determine the best method from training cohort (described in section "Material and Methods") to divide patients into luminal/basal groups, then further validated in the other cohort. Thus, the training and validation cohort make sense in this research.

Comments 2. Do the author have any sequencing/expression data from the cohort for molecular classification which can correlate the IHC classification?

Reply 2. The tumor tissues were obtained from department of pathology in the form of FFPE, which were not suitable for RNA-seq, so there were no sequencing/expression data. However, to support the IHC classification, we further analyzed mRNA and protein expression data from TCGA database (Figure 4). Patients were classified into luminal and basal subtypes according to the criteria described in manuscript. Survival curve of YAP_pS127 and Cyclin B1 of these patients showed that higher levels of either YAP_pS127 or Cyclin B1 indicate tendency of better prognosis, although not statistically significant (Figure 4A&B). While the combination of Cyclin B1 and YAP_pS127 defined a group of patients with significant good survival (Figure 4C), which is consistent with IHC result. Although the mRNA and protein level correlated well (Figure 4D&E), we did not find the same trend at mRNA level (Figure S7C&D).

Changes in the text. Details were described in section "Material and Methods", "Results", "Figure 4" and "Figure S7".

Minor comments:

Comments 3. Line 175, "external validation cohort" is a separate cohort or the same cohort applied for the rest of the study? If it is the same, the "external" should be removed.

Reply 3. The external validation cohort is the same cohort for the rest of the study and the word "external" were removed.

Changes in the text. "External" were removed.

Comments 4. What are the detailed cut-offs for CCNB1 and YAP1? How could these cut-offs be applied in clinical practice?

Reply 4. The cut-offs for CCNB1 and YAP1 were 100 and 160, respectively. That is, if the IHC score (intensity × percentage) > cut-off, the tumor were defined as high-expression group. The risk stratification system can be adopted in clinical practice similar to PD-L1 or HER-2, which requires evaluation of both intensity and percentage of stained cells. Further, examples of different staining intensity were shown in supplementary data (Figure S6), intensity was recorded as followed: 0 for no staining, 1 for weak staining, 2 for median staining and 3 for strong staining.

Reviewer B

The authors retrospectively analyzed molecular subtyping in bladder cancer using immunohistochemistry. Bladder cancer could be divided into basal and luminal subtypes using CK5/6, CK14, and GATA3. Luminal subtypes showed worse survival. When using YAP1 and CCNB1 to further stratify luminal subtypes, distinct prognostic subtypes were identified.
The following topics need to be addressed:

Comments1. How were the training and validation cohort selected? By year or randomly?
Reply 1. The training and validation cohorts were selected mainly by year, patients underwent surgery between 2006-2014 were assigned to training cohort while those between 2011-2016 but not included in training cohort were assigned to validation cohort. The baseline characteristic of the 2 cohort showed no statistical significance (Table 1).
Changes in the text. The selection method of training and validation cohort were detailed in section "Material and Methods-Patients".

Comments 2. It is very hard to read the Kaplan Meier curves in the Figures because they are very blurry
Reply 2. The Kaplan-Meier curves were replaced with high-resolution version (see file "Figure panel.pptx" and "supplementary.pptx").

Comments 3. You found no association of subtypes with survival in NMIBC, which is in contrast to other studies using KRT5 and KRT20 mRNA expression. You should discuss your results with the following studies:
Breyer et al., Virchows Arch 2017, PMID: 28074276
Sikic D et al., Life (Basel) 2021, PMID: 34209360
The reason for the differences could be because of the different methods (immunohistochemistry vs PCR) but also because your smaller cohort of patients with NMIBC
Reply 3. Thanks to this important suggestion, we have discussed the difference regarding the prognosis of NMIBC subtypes in this study and previous reports. The difference may be caused by the following reasons. ① NMIBC patients selected in our study included Ta patients (35.7%) while the study mentioned by reviewer B included mainly T1 high-risk patients, which had worse prognosis. ② The end point was defined as cancer-specific survival in our study while other study chose recurrence or progression as the endpoint. Due to the 5-year survival

of 70%-96% among NMIBC (PMID: 32183076), few patients met the endpoint defined in our study and showed no statistical significance. ③ The difference in detecting methods mentioned by reviewer B.

Changes in the text. The cause of this difference was detailed in line 276-280.

Comments 4. The passage in line 256-264 should be moved to line 244.

Reply 4. The manuscript was revised as the reviewer required.

Changes in the text. Passage was moved as required.

Comments 5. An English native speaker should check the manuscript.

Reply 5. The manuscript was checked by English native speakers and revised.