Peer Review File

Article information: https://dx.doi.org/10.21037/gs-21-678

Reviewer A

Well-written and concise study in which TCV-PTCs are compared to classic/conventional PTCs with focal TCV features (<30% of cells). Large cohort (>500 cases). TCV-PTCs were overrepresented in terms of tumor size, ETE and lateral LN mets, but patient oucomes in terms of death were comparable. Based on this, the authors propose a change in the ATA recommendations to move CPTC with focal TCV features to the intermediate group. I have a few suggestions for the authors to consider:

Comment 1. The conclusions are too strong. In this study, TCV-PTC was overrepresented in terms of several clinical risk features compared to conventional PTCs with focal TCV features only, but the authors advocate that the patient outcome was similar for both groups - which is based on very few samples (only 5.7% and 3.1% of the patients of each tumor group exhibited recurrences). To change the ATA recommendations, multi-center studies should be performed (which the authors touch upon in their discussion). Also, this relates to my second query below.

Reply 1: I agree that the conclusions are too strong. And, we need multi-center studies to prove our results. However, TVCPTC and classic PTC with TCF have similar clinical features and oncologic outcomes in our study. Thus, we are cautious, but here's a suggestion for a "potential" re-classification of classic PTC with TCF from low to intermediate risk in the ATA risk stratification system.

Comment 2. The reason for the low frequencies in terms of recurrences must be discussed. I assume the short follow-up time is one aspect, but please comment on the tumor cohort in terms of staging groups. The T stage in Table 2 is not subclassified as pT1a or b, which should be provided. I suspect most of the tumors in this cohort are in fact papillary microcarcinomas (pT1a). If so, the authors should provide a reasonable explanation as to why they chose to call sub-cm PTCs as TCV-PTCs instead of microcarcinomas (as per WHO 2017). Please also explain if their results are comparable with previous studies, if this work is mainly focused on pT1a tumors? Could the large pT1 group explain the general favourable outcomes? And could the results have been different if only pT1b tumors and above were included?

Reply 2: Thank you for your kind considerations. First of all, T1a in each group was 76 (76/113, 67.3%) and 200 (200/296, 67.6%) patients, respectively. Actually, most of the tumors in our cohort was not PTMCs. The tumor size of our cohort was 1.2cm. For this reason, we did not separate T1a and T1b. PTMC is a classification only for size. However, there are pathologically various variants of PTC. One of the variants of PTC is TCVPTC. PTC was classified was TCV if it contained at least 30% tall cells in the entire tumor volume without tumor necrosis or significant mitotic activity. So, we call TCVPTC pathologically.

T1a accounts for about 50% of total cohort. This work is not mainly focused on pT1a tumors, but all T stages. Even if we included only T1b and above tumors, the results

would not be significantly different.

Comment 3. The study could benefit from the inclusion of a group of conventional/classical PTCs for comparison with the two existing groups as a "control", but this may fall outside the scope of this article. It should be mentioned in the Discussion though.

Reply 3: Thank you for your kind comments. I added about classic PTCs in the limitation section.

Changes in the text: Third, we did not include classic PTC in this study. Several studies have already confirmed that TCVPTC or classic PTC with TCF is more aggressive than classic PTC. (see Page 17, line 4-5)

Comment 4. The high percentage of cases with BRAF V600E mutation is probably the reason for the aggressive features associated to the TCV cells. However, the authors should also mention the role of synchronous TERT promoter mutations and TERT expression in this aspect. Indeed, TERT promoter mutational testing (or even TERT gene expressional analyses using qRT-PCR or in situ hybridization) could be a prognostic force to be reckoned with in the near future, which may complement or even enhance the prognostication of thyroid tumors. I would suggest the authors to discuss this aspect and also cite relevant papers dealing with these aspects of risk assessment in well-differentiated thyroid cancer (for example; PMID: 24476079, PMID: 23766237, PMID: 31286848 and PMID: 34011619). Indeed, future assessment of these tumors may demand histological as well as molecular work-up.

Reply 4: I totally agree with you. BRAF mutation is very important for the prognosis of thyroid cancer, especially TCVPTC. Recently, TERT promoter mutations and TERT expression could be important prognostic factor. It is a well-known fact that the presence of TERT mutation has a poor prognosis. However, we did not include TERT mutation in this study. Because TERT mutation test has been started in 2018 in our hospital. Most of the patients included in this study did not undergo the TERT promoter mutational testing. We will definitely include TERT test in our next study.

Reviewer B

Its an interesting study but confirms prior studies with similar tendencies for increased aggressiveness of TC features.

Comment 1. The two groups do not appear to be adequately matched and therefore improved matching is warranted for this comparison.

Reply 1: To reduce the effect of selection bias and potential confounding factors, propensity score matching is one of good statistical method. However, propensity score matching has a disadvantage in data loss. It is a good method if the cohort difference between the two groups is large, but it is not good if the cohort difference is not large. There is not significant difference in cohort size between TCVPTC and classic PTC with TCF groups. Thus, we did not use propensity score matching method in this study.

Comment 2. The authors need a control group of Classic PTC.

Reply 2: Thank you for your kind comments. I added about classic PTCs in the limitation section.

Changes in the text: Third, we did not include classic PTC in this study. Several studies have already confirmed that TCVPTC or classic PTC with TCF is more aggressive than classic PTC. (see Page 17, line 4-5)

Comment 3. The authors need to adhere to the ATA 50% cut off I think to establish their point. I realized The WHO criteria but if the point to to reestablish the ATA guidelines this needs to be the standard.

Reply 3: There is a discrepancy in the definition of TCVPTC between ATA and WHO classification. According to the ATA guidelines, TCV is characterized by predominance (>50%) of tall cells. However, WHO classification adopted the cutoff of 30% of tall cells in tumor cells. There is no clear consensus on the cutoff of tall cells for defining TCVPTC. Recently, there has been a tendency to use lower thresholds for defining TCVPTC. Many studies have already demonstrated the definition of TCVPTC as the cutoff of 30%. In our hospital, we define the TCVPTC according to the WHO classification.