Peer Review File

Article information: https://dx.doi.org/10.21037/cco-23-89

Reviewer A

Stepwise diagnostic modalities of the preoperative examination of thyroid nodules evolve with technical advancements in hospitals, including ultrasonography, thyroid function serology, fine-needle aspiration (FNA) cytopathology and molecular testing. In this study, Yang et al. presented a stepwise analysis of thyroid diagnostic modalities with genomic imprinting detection in the malignancy risk stratification of thyroid nodules. The inclusion of QCIGISH test enhanced the existing model's diagnostic performance (AUROC: 0.95; 95% CI 0.91 to 1.00, P=0.007). The study was well designed, the manuscript was structurally well developed, and the report was clearly written and informative. In particular, authors demonstrated an interpretation of molecular imprinting detection findings along with ultrasonographic, serological, cytopathologic results using 4 representative cases, providing clinicians or physicians an example how to read molecular testing results to reach a more accurate and definitive preoperative evaluation especially for morphologically indeterminate thyroid nodules and cases with potentially discordant results among standard modalities.
In addition, there are some concerns that need to be addressed.

Comment 1: In the Highlight Box, what is known and what is new remain unclear and need to be rewritten.
Reply 1: We appreciate the reviewer(s) for giving us the opportunity to clarify certain statements we've presented in the manuscript. We have rewritten the narrative at this portion of the **Highlight Box** to emphasize the new contributions of our current study with respect to existing research.
Changes in the text: We have reflected our changes on **Page 3**, **Line 10-14** of the **Manuscript**.

Comment 2: In Method, the paragraph "Sample size calculation …" should be moved up to the end of the section of Patients.
Reply 2: We transferred the paragraph describing the details of the sample size computation from the **Statistical Analysis** to **Patients** section under **Methods**, as advised.
Changes in the text: We have reflected our changes on **Page 10**, **Line 12-22** of the **Manuscript**.

Comment 3: According to the diagram Figure 1, it looks like QCIGISH detection and Cytopathology examination were performed side-by-side. Was QCIGISH detection run in those FNAs under Bethesda II to V categories or in all 326 FNA samples? Please clarify and modify the diagram. How many samples were used in developing model A and how many in developing model B?

Reply 3: We thank the reviewer(s) for highlighting this valid concern and allowing us to carefully explain our study design and workflow diagram.
The reviewer was correct to assume that QCIGISH detection and cytopathology examination were simultaneously conducted in a blinded manner. In effect, QCIGISH was in fact performed for 320 of the 326 prospective cases with 6 excluded due to poor RNA quality or inadequate cell quantity prior to QCIGISH testing. From this number, the subsequent modeling process was only conducted for 114 Bethesda II to V cases with complete preoperative diagnostic and postsurgical confirmatory data for analysis. The ultrasound examination and thyroid function serology information for these 114 cases were then retrospectively gathered. Base model A and stepwise models B, C and D – all under Figure 1 were therefore developed using the same set of 114 cases with complete ultrasonography, thyroid function serology, FNA cytopathology and QCIGISH molecular imprinting diagnoses.
We apologize for the confusion and have simplified our study design and workflow diagram accordingly as recommended. We have also added a description in the manuscript to aid in clarifying this concern.
Changes in the text: We have reflected our changes on **Figure 1**. Additionally, we have added narratives reflecting the aforementioned changes on **Page 13**, **Line 3-7** of the **Manuscript**.

Comment 4: The data of QCIGISH testing could be presented with more details to show the feasibility of more widespread adoption of this test, such as sensitivity, specificity, PPV and NPV in this particular study, though these have been discussed in the Discussion section.
Reply 4: The reviewer(s) highlighted a valid point about how providing more details regarding the diagnostic performance of QCIGISH will demonstrate its feasibility for clinical application. To this end, we have included an additional portion in the manuscript detailing the malignancy rates across all the factors studied. Particularly for QCIGISH, we have provided the estimated sensitivity, specificity, NPV and PPV as advised.
Changes in the text: We have reflected our changes on **Page 14**, **Line 1-16** of the **Manuscript**.

Comment 5: The QCIGISH detection results will fall into five grades (Grades 0, I, II, III and IV), with Grades 0 and I considered as negative and Grades II, III and IV considered as positive. The rationale of classification into negative or positive was not clearly described. The reviewer wonders whether QCIGISH testing result from FNA samples can be confirmed in the follow-up surgical tumor tissues.
Reply 5: We agree with the comments made by the reviewer(s) about potentially failing to provide more details regarding the QCIGISH grading process. We have added a narrative in the **QCIGISH Detection** portion of the **Supplement Methods** discussing what each individual QCIGISH grade represents and the motivation for determining the threshold for differentiating between the negative and positive categories. A previous

study was cited which provides more details regarding the QCIGISH grading model development process.
As a confirmation of the clarification made by the reviewer, QCIGISH could be performed on the follow-up post-surgical tumor tissues (or any histologic specimen in general). However, this was not conducted in the context of this study as post-surgical histopathology served as the gold standard from where all formulated ROC curves which plotted the true positive and false positive rates, and subsequently AUROC estimates were derived from. This is in addition to the overall clinical objective of the QCIGISH test which is to serve as an adjunctive test for the pre-surgical diagnosis of thyroid nodules in the form of cytology specimens.
Changes in the text: We have reflected our changes on **Page 3**, **Line 16-22** and **Page 4**, **Line 1-2** of the **Supplement Methods**.

Comment 6: Table 3 was not shown or mentioned in the text. Explain the malignancy odds of individual diagnostic modalities for thyroid nodules. Multivariate analysis of ultrasonographic, serological, or cytopathologic assessment alone showed no significant OR in malignancy stratification of thyroid nodules. Describe how the univariate and multivariate analyses were done and their differences in the results.
Reply 6: We sincerely apologize to the reviewer(s) for this oversight. We were supposed to cite Table 3 in the narrative discussing the outcome of the multivariate logistic regression model. Driven by these recommendations, we now included as part of the **Results** section, the respective odds ratios obtained from the univariate and multivariate analyses. We would like to clarify however, that due to the stepwise nature of the modeling process where a new modality is sequentially assessed by addition to the previously combined modalities, we feel it's more appropriate to update the Table 3 multivariate analysis results to accurately reflect how the odds ratios change in a stepwise manner.
Changes in the text: We have reflected our changes on **Page 15**, **Line 4-13** and **Line 17-21** of the **Manuscript**, as well as on **Table 3**.

Comment 7: Cite literatures with more cautious. For example, obviously there is a question to "BRAF V600E mutation can only be detected in up to 83% of papillary thyroid cancers,..." in line 393. Actually, "BRAF V600E mutation occurring in 40–45% of PTC" were stated in the two articles cited as #22 (https://pubmed.ncbi.nlm.nih.gov/15947100/) and #51 (https://www.nature.com/articles/nrdp201577). This number can vary a lot in some cohorts with unreasonable sample sizes. It also changes reasonably with the advances in detection methods. A most recent study showed BRAF V600E variant was identified 53.0% (158 /298) of PTC tumors by ultrasensitive digital PCR assays (https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2807711). References need to be updated.
Reply 7: We truly appreciate these valid concerns highlighted by the reviewer(s) and likewise seek for their understanding on this oversight. We have indicated the correct reference publication where the declared figure was derived from. In addition, we have also included the references suggested by the reviewer(s). Based on the important comments made, we have also revised portions of the **Discussion** by specifying that the

presented numbers were conditioned by various factors including the cohort characteristics, sample size, and detection methods used.
Changes in the text: We have reflected our changes on **Page 19**, **Line 20-22** and **Page 20, Line 1-2** of the **Manuscript**.

Reviewer B

1. Please provide the full name of "AUROC" "CI" in the abstract and "miRNA" "lncRNAs" "ROC" "CI" "PPV" "NPV" "OR" "ISH" in the main text. Please also check through your article to make sure **all** the abbreviated terms have been defined when they **FIRST** appear in the Abstract and the main text.

   Response: We have provided the full names of the abbreviations in both the abstract and the main text. We have additionally checked all the abbreviated terms and provided their full names upon first appearance.

2. Please provide the full names of the abbreviated terms in the highlight box.

   Response: We have provided the full names of the abbreviations used in the highlight box.

3. Please check if any reference should be added since you mention "studies".

   "Although, certain studies have shown that increased FT4/FT3 ratio could be a more reliable malignancy predictor for thyroid nodules(17)."

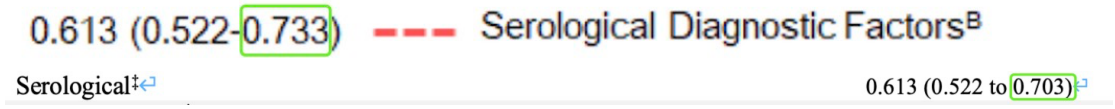
   Response: We have corrected the oversight and added references 18 and 19 accordingly.

4. Figure and tables
   - **All abbreviations** in figures/tables and legends should be explained. "QCIGISH" in Figure 1, "ROC" "AUROC" "AUC" "CI" "ACR TI-RADS" "FT4/FT3" "QCIGISH" in Figure 2, and "IQR" in Table 1 for example. Please check all abbreviations and provide the full names in the corresponding figure legend/table foot.

     Response: We have provided the full names of all the abbreviations in the figure legend.

   - Do not insert figure legends in the Figures. Please remove them and resubmit Figure 2, 3 in jpg/tiff format.
     Response: We have removed the figure legends for Figures 2 and 3. We have also re-attached the revised versions of Figures 2 and 3 in tiff format.

- Please recheck the data in Figure 2.



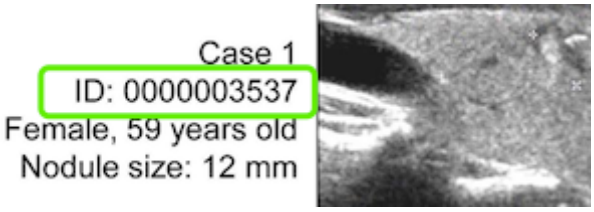
"the serological (AUROC: 0.61; 95% CI 0.52 to 0.70; optimism-adjusted AUROC: 0.60) factor (P<0.001 to P=0.003) (Tables 2 and S1, Figure 2)."

Response: The correct value for the highlighted number was '0.703'. We have made corrections to Figure 2B to reflect the changes. No revisions were however necessary in the manuscript text.

- Figure 3: please indicate the staining method for cell maps in the legend.

Response: We have indicated the staining method for the cytopathology images.

- Figure 3: please remove or mask any information which may reveal the patient's privacy.



Response: We thank the editors for this recommendation. We have removed the patient IDs as advised.

- The unit is repeated in Table 1, 2. Please remove.

Sex (%)	
Male (%)	33 (29%)
Female (%)	81 (71%)

Response: We have removed the indicator '%' in Tables 1 and 2, as highlighted.

- It is suggested to recheck the ratio.

"The median age at diagnosis was 50 years (IQR: 36 - 62) with female predominance (71.1% versus 28.9%; female to male ratio, 3:1)."

Response: We have updated the female to male ratio from '3:1' to '2.5:1' to reflect a more accurate representation of the aforementioned metric.

- Please recheck the data in Table 1, whether "5" would be more appropriate after rounding.

Adenomatous Goiter (%)	6 (6%)

Response: We have checked the data and corrected the percentage from '6%' to '5%'.

- Please recheck the data in Table 2.

Low (%)	54 (44%)
High (%)	60 (56%)

Response: We have checked the data and corrected the percentages indicated as '44%' and '56%' to '47%' and '53%', respectively.

- Please recheck the data in the following sentence.

  "Increasing proportions of histopathologically malignant cases were similarly observed for the combined Bethesda categories II, III and IV at 33.28% (24/70) and Bethesda category V at 90.91% (40/44)."

  Response: We have checked the data and corrected the percentage from '33.28%' to '34.28%'.

- Please recheck the explanation of "‡" in Table 2 footnote.

  Response: We have checked the explanations labeled as "‡" under the Table 2 footnote and removed the redundant items.

- Table 3: the heads are same. Are there any differences between them? It's not allowed to have the same heads in a Table.

Univariate Analysis		Stepwise Multivariate Analysis							
(*n*=114)		(*n*=114)							
OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†

Response: We have indicated unique and separate labels for the table heads as recommended. We have also provided additional footnotes to detail the differences in the updated table.

5. Supplementary

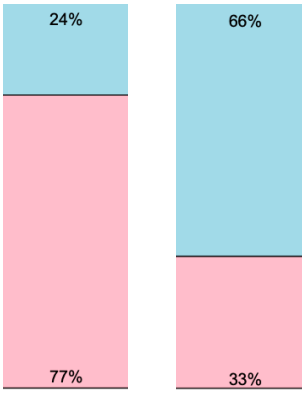
- **All abbreviations** in supplementary figures/tables and legends should be explained. "QCIGISH" in Figure S1, and "CI" "ROC" "ACR TI-RADS" "TSH" "FT3" "FT4" in Table S1 for example. Please check all abbreviations and provide the full names in the corresponding figure legend/table foot.

  Response: We have provided all the necessary abbreviations in the supplementary figures and table, as highlighted.

- Please indicate the staining method in Figure S1 legend.

  Response: We have indicated "QCIGISH staining" under Figure S1 legend.

- It is suggested to recheck the data in Figure S3.



Response: We have checked the data and corrected the percentages accordingly.

- Please provide a table header for the first column in Table S1.

  Response: We have provided an appropriate table header for the first column of Table S1, as recommended.

- In the Supplementary Materials and Methods, please provide the full name of the abbreviated terms when they first appear.

  Response: We have provided the full name of the abbreviations used in the Supplementary Materials and Methods upon first appearance.

- In the Supplementary Materials and Methods, the citations of the references should be in **regular round brackets with a space before**.

  Response: We have revised the citations following the requested format.

6. The publication information of ***Ref 34*** seems incomplete, please supplement the page number/doi link for it.

   Response: We have added the necessary volume and page number details for Ref 34.

7. Ethical statement
   - The Declaration of Helsinki should be described in the **Methods.** Please add.
   - Please also indicate the status of the ethical approval and informed consent in the Methods section.

   Response: We have added the Declaration of Helsinki and the status of the ethical approval and informed consent under the Methods section.