

RESPONSE TO REVIEWERS COMMENTS, WJR

Our response to each of the reviewers comments is provided below each of the queries and is highlighted for ease of viewing. All queries are addressed.

Reviewer #1:

Scientific Quality: Grade B (Very good)

Language Quality: Grade A (Priority publishing)

Conclusion: Minor revision

Specific Comments to Authors: This a very interesting manuscript describing inter-reader agreement of radiology residents in using the ACR-RADDS system. It is well-written, but there remain some problems (see below). Major points

a) Materials and Methods: “The type of ultrasound make, model, or platform were not considered in the selection process.” I cannot understand this way, because US image is largely influenced by these factors (frequency, degree of compound or smoothing, --).

This was a retrospective study so we did not have control of the source data. As a tertiary referral centre, the primary US examinations were sourced from a variety of different referral clinics and hospitals. As such, we were not able to standardize for different ultrasound makes/models/platforms.

However, it is important to be clear that ACR TI-RADS is a robust internationally adopted US risk stratification system and is designed to be usable and applicable across the different US vendor platforms and technologies.

We would also like to reassure the reviewer that all ultrasound platforms were in compliance with stringent Health Canada approved diagnostic quality standards as required by the regulatory board for diagnostic medical equipment. Examinations were performed by diagnostic ultrasound technologists that were fully trained and certified in performing thyroid US, and licenced by our provincial board for Medical Sonography.

Finally, and as a safeguard, all US examinations were pre-screened by the research steering committee to ensure that these met stringent technical and diagnostic quality standards for use in the project. As already mentioned in the Material and Methods section of the manuscript,

“Nodules with non-diagnostic image quality, incomplete nodule visualization, and absence of a cine clip covering the entirety of the nodule were excluded”.

b) Please mention how to teach thyroid US in the training in detail.

The following has been added for clarification in the manuscript:

“Four weeks after the readers had completed the pre-training assessment; a one hour-long teaching session including a PowerPoint presentation illustrating important features of ACR TI-RADS was provided to the readers along with a word document summarizing common areas of disagreement in nodule characterization. The teaching session went through in a step-by-step fashion the 5 main sonographic

features used for nodule scoring in ACR TI-RADS: 1. Composition 2. Echogenicity 3. Shape 4. Margin and 5. Echogenic foci. Each feature and how it should be appropriately interpreted was discussed and illustrated by examples. The readers were given ample opportunity to ask questions, and the consensus panel provided focused clarification to readers in areas of reader uncertainty”.

c) Please show some representative US images (good agreement, poor agreement).

We have now added Figures 1 and 2 to the manuscript.

Figure 1 shows a thyroid nodule where all 3 readers showed perfect concordance and classified the nodule as TI-RADS level 5. Figure 2 shows a thyroid nodule where there is discordance between readers – 2 of the readers scored the nodule as TI-RADS level 4 while the 3rd reader scored this as TI-RADS level 5.

Minor points a) Clinical experience: three fellowship-trained staff radiologists with between 1 and 14 years. In my opinion, 1 year of clinical experience is insufficient.

We respectfully disagree with the reviewer opinion.

The radiologist with 1 year clinical experience has undergone advanced subspecialty training in ultrasound including on ACR TI-RADS, and is board certified by the American College of Radiology and the Royal College of Physicians and Surgeons of Canada, in Diagnostic Radiology. Furthermore, that particular radiologist was selected to be on the consensus panel because he is regarded as one of the finest diagnostic radiologists in our institution.

b) Material and Methods: What kind of disagreement was resolved by re-review?

Disagreement on *scoring of the ACR TI-RADS level* was resolved by re-review.

The following statement has been added:

“Any disagreement on the scoring of nodules for the ACR TI-RADS level was resolved by re-review and consensus discussion”

Reviewer #2:

Scientific Quality: Grade A (Excellent)

Language Quality: Grade A (Priority publishing)

Conclusion: Accept (High priority)

Specific Comments to Authors: This is an excellent manuscript with an important topic, sound analysis and solid writing. Especially, you addressed very well the strengths and limitations of your study in Discussion.

We thank Reviewer 2 for the excellent appraisal of our manuscript. We wholeheartedly agree with the reviewers comments.

Reviewer #3:

Scientific Quality: Grade A (Excellent)

Language Quality: Grade A (Priority publishing)

Conclusion: Accept (General priority)

Specific Comments to Authors: This study demonstrates a statistical significant improvement in inter- reader agreement among radiology residents before and after a teaching session of ACR TIRADS compared to expert consensus. The study therefore proposes the use of dedicated training for residents as an affective mean to improve capability in interpreting thyroid nodule ultrasound . Further studies should take place comparing diagnostic improvement to pathological reference and not only expert consensus as was performed in this study.

We thank the reviewer for comments.

We have already acknowledged the lack of a pathological gold standard in the limitations paragraph of the Discussion. As mentioned in our text, “this study is designed primarily to evaluate inter-reader agreement and as such as an expert consensus panel is a reasonable reference standard, and one that simulates ‘real world’ clinical practice”.

Furthermore, other published studies have also used an expert consensus panel rather than pathology as a reference standard in cases where inter-reader agreement is the primary focus of analysis, including:

Pi Y, Wilson MP, Katlariwala P, et al. Diagnostic accuracy and inter-observer reliability of the O-RADS scoring system among staff radiologists in a North American academic clinical setting. *Abdominal radiology (New York)* Oct 2021;46(10):4967-4973

4 LANGUAGE POLISHING REQUIREMENTS FOR REVISED MANUSCRIPTS SUBMITTED BY AUTHORS WHO ARE NON-NATIVE SPEAKERS OF ENGLISH

As the revision process results in changes to the content of the manuscript, language problems may exist in the revised manuscript. Thus, it is necessary to perform further language polishing that will ensure all grammatical, syntactical, formatting and other related errors be resolved, so that the revised manuscript will meet the publication requirement (Grade A).

Authors are requested to send their revised manuscript to a professional English language editing company or a native English-speaking expert to polish the manuscript further. When the authors submit the subsequent polished manuscript to us, they must provide a new language certificate along with the manuscript.

Once this step is completed, the manuscript will be quickly accepted and published online. Please visit the following website for the professional English language editing companies we recommend: <https://www.wjgnet.com/bpg/gerinfo/240>.

The senior author is an experienced scientific writer with over 110 peer reviewed PUBMED indexed publications in English speaking journals. He can guarantee the quality of the material and that it is written professionally using the highest standards of English writing. All authors are Canadian and English is our first language.

5 ABBREVIATIONS

In general, do not use non-standard abbreviations, unless they appear at least two times in the text preceding the first usage/definition. Certain commonly used abbreviations, such as DNA, RNA, HIV, LD50, PCR, HBV, ECG, WBC, RBC, CT, ESR, CSF, IgG, ELISA, PBS, ATP, EDTA, and mAb, do not need to be defined and can be used directly.

The basic rules on abbreviations are provided here:

(1) Title: Abbreviations are not permitted. Please spell out any abbreviation in the title.

done

(2) Running title: Abbreviations are permitted. Also, please shorten the running title to no more than 6 words.

done

(3) Abstract: Abbreviations must be defined upon first appearance in the Abstract.
Example 1: Hepatocellular carcinoma (HCC). Example 2: *Helicobacter pylori* (*H. pylori*).

done

(4) Key Words: Abbreviations must be defined upon first appearance in the Key Words.

done

(5) Core Tip: Abbreviations must be defined upon first appearance in the Core Tip.
Example 1: Hepatocellular carcinoma (HCC). Example 2: *Helicobacter pylori* (*H. pylori*)

done

(6) Main Text: Abbreviations must be defined upon first appearance in the Main Text.
Example 1: Hepatocellular carcinoma (HCC). Example 2: *Helicobacter pylori* (*H. pylori*)

done

(7) **Article Highlights:** Abbreviations must be defined upon first appearance in the Article Highlights. Example 1: Hepatocellular carcinoma (HCC).

Example 2: *Helicobacter pylori* (*H. pylori*)

done

(8) **Figures:** Abbreviations are not allowed in the Figure title. For the Figure Legend text, abbreviations are allowed but must be defined upon first appearance in the text.

Example 1: A: Hepatocellular carcinoma (HCC) biopsy sample; B: HCC-adjacent tissue sample. For any abbreviation that appears in the Figure itself but is not included in the Figure Legend textual description, it will be defined (separated by semicolons) at the end of the figure legend. Example 2: BMI: Body mass index; US: Ultrasound.

done

(9) **Tables:** Abbreviations are not allowed in the Table title. For the Table itself, please verify all abbreviations used in tables are defined (separated by semicolons) directly underneath the table. Example 1: BMI: Body mass index; US: Ultrasound.

done

6 EDITORIAL OFFICE'S COMMENTS

Authors must revise the manuscript according to the Editorial Office's comments and suggestions, which are listed below:

(1) *Science editor:*

This manuscript assessed the inter-reader reliability of TI-RADS in radiology trainees, and demonstrate the learnability of the ACR TI-RADS system. It is recommended that comparing diagnostic improvement to pathological reference; type of ultrasound make, model, or platform will affect the learning effect, please explain how to ignore it.

Language Quality: Grade A (Priority publishing)

Scientific Quality: Grade B (Very good)

1. Lack of Pathological reference standard

We have already acknowledged the lack of a pathological gold standard in the limitations paragraph of the Discussion. As mentioned in our text, "this study is designed primarily to evaluate inter-reader agreement and as such as an expert consensus panel is a reasonable reference standard, and one that simulates 'real world' clinical practice".

Furthermore, other published studies have also used an expert consensus panel rather than pathology as a reference standard in cases where inter-reader agreement is the primary focus of analysis, including:

Pi Y, Wilson MP, Katlariwala P, et al. Diagnostic accuracy and inter-observer reliability of the O-RADS scoring system among staff radiologists in a North American academic clinical setting. *Abdominal radiology (New York)* Oct 2021;46(10):4967-4973

2. Type of US make, model and platform on learning effect

This was a retrospective study so we did not have control of the source data. As a tertiary referral centre, the primary US examinations were sourced from a variety of different referral clinics and hospitals. As such, we were not able to standardize for different ultrasound makes/models/platforms.

However, it is important to be clear that ACR TI-RADS is a robust internationally adopted US risk stratification system and is designed to be usable and applicable across the different US vendor platforms and technologies. As such, we don't not believe that the 'type of US make, model and platform' will be a significant differentiator in the learnability of ACR TI-RADS – the ACR committee designed it to be universal across US platforms, makes and models.

We would also like to reassure the Science Editor that all ultrasound platforms were in compliance with stringent Health Canada approved diagnostic quality standards as required by the regulatory board for diagnostic medical equipment. Examinations were performed by diagnostic ultrasound technologists that were fully trained and certified in performing thyroid US, and licenced by our provincial board for Medical Sonography.

Finally, and as a safeguard, all US examinations were pre-screened by the research steering committee to ensure that these met stringent technical and diagnostic quality standards for use in the project. As already mentioned in the Material and Methods section of the manuscript,

"Nodules with non-diagnostic image quality, incomplete nodule visualization, and absence of a cine clip covering the entirety of the nodule were excluded".

(2) Company editor-in-chief:

I have reviewed the Peer-Review Report, full text of the manuscript, and the relevant ethics documents, all of which have met the basic publishing requirements of the World Journal of Radiology, and the manuscript is conditionally accepted. I have sent the manuscript to the author(s) for its revision according to the Peer-Review Report, Editorial Office's comments and the Criteria for Manuscript Revision by Authors. Please provide the original figure documents. Please prepare and arrange the figures using PowerPoint to ensure that all graphs or arrows or text portions can be reprocessed by the editor. Authors are required to provide standard three-line tables, that is, only the top line, bottom line, and column line are displayed, while other table lines are hidden. The contents of each cell in the table should conform to the editing specifications, and the lines of each row or column of the table should be aligned. Do

not use carriage returns or spaces to replace lines or vertical lines and do not segment cell content.

The tables have been modified following specifications above.

The Figures (Figures 1 & 2) are provided both in individual TIFF formats and in a powerpoint file. We did not need to introduce annotations on powerpoint as the nodules were already annotated on the source images on PACS.