Peer Review File
Article information: https://dx.doi.org/10.21037/jtd-23-998

Reviewer A

Overall, this is a very well-written paper that is particularly revealing when it comes to the lack of standardization in the field of interventional pulmonology.

Highlight box:
1. Missing the word "yield" in the second part of the first sentence.

The authors do an excellent job highlighting dramatic inter-rater variability. Perhaps equally important is the accuracy of the interpretation. Are the authors able to provide any sense of how often the study subjects were correct in their classification? Also, the authors may consider proposing a standardized definition of the terms concentric, eccentric, and no image, so that future studies may benefit from improved consistency in nomenclature. This is especially true as data begin to emerge from the new REBUS probe with integrated biopsy needle (iNod by Boston Scienitific).

Neither "concentric" nor "eccentric" should be capitalized unless it is the first word in a sentence.

Reply:

1. The authors thank this reviewer for their comments. The objective of the study was to describe and demonstrate the current subjectivity within the field of diagnostic bronchoscopy for the biopsy of peripheral pulmonary lesions. While there is no formal definition or standardized criteria, currently lesions are described as concentric when obtained images show the radial probe within and surrounded by the lesion. A lesion is described as eccentric when the images obtained show the probe towards one side of the lesion, suggesting the lesion is adjacent to the airway.
   As there is no formal standardized criteria, we sought not to compare the subjects responses to any "correct" response. This is because any of the subject's responses would only be compared against the original proceduralists interpretation, which is also subjective. Notably, a goal of this study is to set the groundwork for a follow-up study proposing a new framework for the interpretation of radial EBUS imaging and to evaluate for improvements in observer variability.

Changes in text:

1. "yield" added to first sentence in highlight box.
2. Decapitalized "concentric" and "eccentric" and "no image" throughout the paper

Reviewer B

This is a well written straightforward study comparing intra- en interobserver variability of

rEBUS still images. The authors show high intra-observer consistency, but low inter-observer agreement.
In part this may be explained by a lack of definition in how to define intralesional positioning, where a slightly off-center position still renders a concentric lesions, as compared to a strict lateral postion.
My biggest concern therefore is twofold:
1. Did the study protocol have a clearly defined definition on how to categorize images, or was this left to the interpretation of the clinician?
2. What is concider the gold standard of concentricity. The authors show the fluoro-image next to the rEBUS image where I would recommend to have 3D-CBCT confirmation of tool-in-lesion confirmation as the gold standard.

Please elaborate on these items in the methods and discussion sections

Reply:

1. The authors appreciate this reviewer's comments for review. The comment by the reviewer actually serves as the initial motivation for this descriptive study. As noted by the reviewer and alluded to in the introduction of our manuscript, currently there does not exist a formal and standardized definition of concentric and eccentric lesions. However, informally, lesions are described as concentric when obtained images show the radial probe within and surrounded by the lesion. A lesion is described as eccentric when the images obtained show the probe towards one side of the lesion, suggesting the lesion is adjacent to the airway. This working definition is one provided in training to all fellows while in training and subjects were asked to grade lesions based on their expert interpretation. This is clarified in the methods section on page 7, line 21-23 as well as the Discussion on page 11, line 1 and 2. This serves to provide the groundwork of proposing a formal definition in a follow-up study.
2. Currently there exists no gold standard for the definition of concentricity for Radial EBUS interpretation and we aimed to establish this point in the introduction of the manuscript. The reviewer offers an intriguing proposal for establishing this definition. An interesting follow-up study would be to publish various EBUS images corresponding to 3D-fluoroscopy or Cone-Beam CT images showing "tool-in-lesion" to further describe the variance seen in Radial-EBUS when this view is obtained.

Changes in text:

1. Methods section, page 7, line 21-23: clarified that there was no formal definition provided to subjects
2. Discussion section, page 11, line 1 and 2: clarified that subjects used their prior training and expert opinion for image interpretation

Reviewer C

The authors provide results of a survey illustrating poor agreement between interventional pulmonologists' interpretations of rEBUS images. The conclusions of the study are important

as discussed by the authors, in particular how misinterpretation of rEBUS may negatively impact clinical practice.

Several aspects of the study should be addressed before publication to strengthen the value of the paper:
Major:
1- One of the key findings of the study was the poor agreement between raters, particularly for rEBUS images with no view (kappa of 0.378 and 0.395 for the two surveys, respectively). As one who is every experienced with rEBUS (and would be akin to a typical survey respondent used in the study), I find this result very unusual. rEBUS images with 'no view' of a lesion are typically the easiest to interpret and would have the least potential for misinterpretation. So it is surprising that these were the ones with poorest agreement between survey respondents. If the authors can clarify these results, and preferably offer an explanation in the discussion, this would be helpful. For example, one possible explanation could be the lack of identifiable lesion borders of large concentric images-- this could confuse an interpreter to rate that image as having no view. Another explanation could be poor image quality; or a combination thereof.

2- It would be helpful to discuss in the Methods section the process and oversight of image selection, quality and distribution, as these would potentially influence interpretation. How were the images chosen for the survey? Were these consecutive cases (regardless of image type and/or quality), randomly selected (again regardless of type or quality), or those selected based on quality and targeted proportional distribution of rEBUS image types? What is the typical field of view/depth used by the authors, and were these standardized during image selection? Were the images distributed digitally and in what color or contrast scheme (it may help to include a sample of the actual images distributed, one each representing concentric, eccentric and no view-- those provided in figure 1 do not specify if these were actual images used in survey).

3- For the 3-month follow up survey, were the images revealed to the interpreters in the same order? Or a random order?

4- The authors should consider discussing in more detail their thoughts on why there is variability of interpretation. Associations with lesion size or location? Is it because concentric/eccentric terms might sometimes be confused with symmetric/asymmetric, which are not synonymous? If the more subtly ('barely') concentric images (ie. asymmetric concentric images) were more likely to be rated eccentric than those that are obviously concentric, that might support such a hypothesis. Whereas if the 'barely' eccentric images were more likely to be rated concentric than obviously eccentric images, it could simply be a matter of image quality and/or 'judgment call.'

5- The authors could propose a standard definition (eg. based on principles of ultrasound physics and airway-lesional relationships). What is the authors' definition of concentric vs eccentric as it applies to rEBUS, and why do they use such definitions?

Minor:

1- In the highlight box, first section, first bullet point, the word 'yield' is missing after 'diagnostic.'

2- In the highlight box, third section, first bullet point sentence is confusing; consider rewording.

3- In the abstract - conclusion, I believe the authors meant to write "strong intra-rater AGREEMENT" rather than "variability."

4- Methods, lines 103-104: consider citing this statement about kappa result interpretation.

5- Discussion, under limitations, line 207 about survey participants being all from same fellowship-- this should also be included in methods section.

6- Were the survey respondents required to complete the survey with all 100 images in one sitting? If so, how might 'image fatigue' impact interpretations toward the end of the survey, and consequently the study results? How was this addressed in the design, if at all?

Reply:

Major

1. We agree with the reviewers' comments regarding the variability of the interpretation of radial EBUS images deemed as "no image". We appreciate and agree with the offered explanations for why this may have occurred and have included this in the manuscript. In hindsight, cases with clear concentric or eccentric images may have been included. As a pragmatic design to the study, we wished to included consecutive cases which resulted in possibly poor-quality images. Due to this, we felt the need to include an option of "no image" to avoid any effect on variability or agreement for the interpretation of concentric or eccentric images if someone felt compelled to choose between only two options when neither was clear.
2. These cases were selected as the most recent 100 cases fulfilling criteria immediately after the study began. Criteria included 100 consecutive cases with available CT scan imaging, radial EBUS images, a finalized pathology report, and one year of follow-up.
3. The respondents were provided the survey with images in similar order. Shuffling of the images was felt to be unnecessary as most images could have been readily identifiable if participants wanted to review their prior responses. For the integrity of the study, subjects were asked to rate the images without reviewing their prior responses.
4. Please see discussion, page 13, line 12-19
5. This was a common comment by several of the reviewers and actually serves as the initial motivation for this descriptive study. As noted by the reviewer and alluded to in the introduction of our manuscript, currently, there does not exist a formal and standardized definition of concentric and eccentric lesions. However, informally, lesions are described as concentric when obtained images show the radial probe within and surrounded by the lesion. A lesion is described as eccentric when the images obtained show the probe towards one side of the lesion, suggesting the lesion is adjacent to the airway. This is the working definition of the authors and is one provided in training to all fellows. Subjects were asked to grade lesions based on their expert interpretation in accordance with their training. Notably, a goal of this study is to set the groundwork for a follow-up study proposing a new framework for the

interpretation of radial EBUS imaging and to evaluate for improvements in observer variability.

Minor:

1. Corrected
2. Added additional comment to highlight box explaining what is known on the subject. This served to add clarification brought up by multiple reviewers as to the current definition of concentric and eccentric views during R-EBUS. Additionally, the original reviewed line 1 in the section under "implications" was deleted as this was felt similar to the line under "key findings" which seemed to better summarize the point being made.
3. Corrected
4. This is information is also included in reference 8 which also provides a reference for the subsequent line. Given the proximity, the reference was noted at the end of these two sentences.
5. As noted on page 8, line 4-5. Clarified that instructions were provided to subjects to not review their prior responses.
6. The subjects were not required to complete the survey in one sitting; however this was not explicitly included in the instructions provided and not addressed in the initial design of the study. While the reviewer raises a salient point regarding the possibility of fatigue effecting effort related to image interpretation, there is no literature to report or guide correction for this phenomenon. Subjects were asked to assess the images to the best of their ability in accordance with their expert and advanced training. Although fatigue and effort may be a factor, the strong agreement seen in the results of the second survey show internal consistency and would suggest effort and fatigue did not play a factor.

Changes in text:

Major

1. See discussion, page 13, line 20-23
2. Added language clarifying these were consecutive cases as well as the fact that digital images from the procedure report were taken for inclusion in the survey. Clarified that Figure 1A and 1B were representative images from the survey.
3. Clarified in methods section, page 8, line 5-6
4. See discussion, page 13, line 12-19

Minor:

1. Corrected as advised
2. See highlight box, section 1, line 1; section 3, line 1
3. Corrected as advised
4. Corrected as advised
5. Corrected as advised, see Methods section, page 7, line 21-22
6. No text added