

## Peer Review File

Article information: <https://dx.doi.org/10.21037/jtd-24-841>

### Reviewer A

This study evaluated the correlation between the AIBA and human experts rating on anatomical and navigational competencies in flexible bronchoscopy. Expert's rating is time consuming and prone to rater bias, thus the AIBA may become a useful alternative to the conventional expert's rating. I have a few comments:

Comment 1. Were there any correlations between the experienced bronchoscopists and novice bronchoscopists on the 4 outcomes of AIBA?

**Reply:** When gathering validity evidence, relationship to other variables by comparing experts to novices is very commonly used. However, we chose not to do so based on the following phrases from the manuscript:

“We chose not to divide the participants into groups based on experience. When collecting validity evidence, it is erroneous to make experienced-novice comparisons for aspect four of Messick's validity framework: i.e., relationship to other variables [35-37]. We therefore chose to gather validity evidence for this aspect by correlating to a gold standard for performance by a validated assessment tool”.

Unfortunately, we are therefore not able to make to comparison as the reviewer requests. In flexible bronchoscopy procedural-volume does not ensure competence, and we therefore chose to follow the recommendations made by CHEST, to assess based on proficiency criteria (mastery learning).

Comment 2. I think skills of scope movement as well as anatomical knowledges are important on the assessment of competence. The AIBA outcomes, especially SP, showed a good correlation not only with expert's anatomy rating but also dexterity-rating, and the authors described bronchoscopists with good anatomical knowledge also showed a higher dexterity level. I think it is true, however, the assessment without a direct measure of dexterity seems to be insufficient. Do the authors think bronchoscope maneuvering skills can be evaluated by AIBA?

**Reply:** We thank the reviewer for highlighting this limitation of AIBA. AIBA is a prototype exploring the first AI assessment and navigation system in bronchoscopy and can currently not be used to assess maneuvering skills. We believe that it will be developed to directly evaluate dexterity, and not just SP which is correlated to it. We have therefore revised the manuscript accordingly.

**Changes in the text:** The AI holds the potential to replace expert raters, if being further developed to provide the bronchoscopists with direct dexterity measures. **This is the first AI navigational system tested in bronchoscopy, and further development of the AI should entail direct dexterity measures as those assessed by the expert raters. In the future, the AI could be implemented in everyday clinical practice to help ensure competent performance before allowing trainees to begin supervised practice on patients.**

## Reviewer B

General:

Comment 1. The authors showcase their study using bronchoscopy AI software to measure bronchoscopy performance among random participants at a European conference. They deconstruct procedural competency into four quantifiable areas that are then analyzed separately.

**Reply:** Please see next reply

Comment 2. Nice job recognizing limitations of the current approach and how elements like SP are best suited to novice bronchoscopists.

**Reply:** We thank the reviewer for his/her comments acknowledging the justification of the study.

Abstract:

1. AIBA not defined prior to first use.

**Reply:** The reviewer is correct, and we have changed the abstract to not use the reference, but just refer to AIBA as the AI.

**Changes in the text:** Line 10 and 14; AIBA → the AI.

Discussion

1. The sentence spanning lines 131-132 seems nonsensical/incomplete. Only in the context of the following sentence does it make sense.

**Reply:** We thank the reviewer for noticing this and we have revised the sentence.

**Changes in the text:** A bronchoscopist should have a high level of anatomy knowledge to ensure inspection of all bronchial segments. Therefore, DC is the first and most widely used metric to assess bronchoscopy competence.

2. Several minor syntax errors, such as line 175 (inversely should be inverse), line 184 (differentiate should be differentiated), line 193 (We encourage who to do a mastery learning?)

**Reply:** We thank the reviewer for his/her thorough examination of the manuscript and the following changes have been made.

**Changes in the text:** Line 175, corrected to inversely, line 184, corrected to differentiated, line 193 corrected to: We encourage fellow researchers to do a mastery learning training study using AIBA.

3. Consider defining mastery learning, since this is a term with which many readers may be unfamiliar.

**Reply:** The reviewer is correct, and we have therefore defined mastery learning when first used.

**Changes in the text:** line 191, This finding indicates that the outcome measure constituting AIBA can be used to set proficiency training criteria, enabling ~~(mastery learning)~~, which is a training modality where trainees practice until these proficiency targets are met (ref inserted).