

## Peer Review File

Article information: <https://dx.doi.org/10.21037/atm-24-171>

### **Comment from Reviewer A**

#### Comment 1:

please add the following paper:

Manual, semiautomated, and fully automated measurement of the aortic annulus for planning of transcatheter aortic valve replacement (TAVR/TAVI): analysis of interchangeability.

Lou J, Obuchowski NA, Krishnaswamy A, Popovic Z, Flamm SD, Kapadia SR, Svensson LG, Bolen MA, Desai MY, Halliburton SS, Tuzcu EM, Schoenhagen P.

J Cardiovasc Comput Tomogr. 2015 Jan-Feb;9(1):42-9. doi: 10.1016/j.jcct.2014.11.003. Epub 2014 Nov 13.

PMID: 25533222 Clinical Trial.

#### Reply 1:

*Thank you for recommending the paper. This study assesses the interchangeability of manual, semiautomated, and fully automated methods for measuring the aortic annulus in patients undergoing transcatheter aortic valve replacement (TAVR). It emphasizes the impact of calcification on measurement accuracy, which further illustrates the challenges associated with aortic root measurement. The findings highlight the importance of manual input in mitigating the effects of calcification.*

#### Changes in the text:

*We have added corresponding statements in the manuscript. (Page 9, Lines: 262-264)*

*Table 3: AI-driven studies on aortic diameter measurement and reference list have been updated with the newly included publication.*

### **Comments from Reviewer B**

#### Comment 1:

why excluding diagnosis-oriented studies or studies without screening?

Reply 1:

*Thank you for your careful review of our manuscript. We appreciate your attention to detail and for identifying the typo. Diagnosis-oriented studies and studies without screening were included in the literature search to evaluate the outcomes of screening and diagnosis generated by the models. We have corrected this error in the revised version of the manuscript.*

Changes in the text:

*We have deleted corresponding statements in the manuscript (Page 4, line: 105)*

Comment 2:

line 102, plain ct is almost never performed in the context of aortic pathology, as it gives limited information on the vessel

Reply 2:

*We fully agree that plain CT is rarely performed alone in diagnosis of aortic aneurysm. However, there is growing evidence that plain CT become a useful tool in addition to contrast-enhanced CT, as our references stated (3,4,6,13) Plain CT scans are utilized to provide aortic diameter assessment and differentiate intramural hematoma (IMH) from other factors of aortic wall thickening.*

Changes in the text:

*We have modified the corresponding statements in the manuscript. (Page 4, lines: 119-121)*

Comment 3:

line 106, in emergency setting, a CTA will almost always be performed, even in CKD patients

Reply 3:

*We fully agree with your insightful comment and we have changed our wordings accordingly.*

Changes in the text:

*We have modified the text as advised. (Page 5, line 133)*

Comment 4:

Consider defining which landmarks are most critical and if AI models currently perform well across all landmarks.

Reply 4:

*Thank you for your constructive comment. Although the landmarks being adopted by literatures vary, most of them include all landmarks suggested by 2010*

*ACCF/AHA/AATS/ACR/ASA/SCAI/SIR/STS guideline. Therefore, we think that all landmarks mentioned in this guideline should be included for standardization. In terms of model performance along these landmarks, please refer to lines 368-373.*

*We have also updated the brief description of landmarks and landing zones according to the guidelines.*

*Changes in the text:*

*We have added the model performance by landmarks in our manuscript. (Page 12, lines 368-373). Also, brief description of landmarks and landing zones have been updated in the section of “Locations and angles”. (Page 6, lines 167-169). Table 2 has been updated accordingly.*

*Comment 5:*

*The use of ICC (Intraclass Correlation Coefficient) to compare measurements is useful; it would be helpful to mention thresholds that indicate acceptable reliability in clinical practice.*

*Reply 5:*

*We agree that it would be helpful to have a pre-defined threshold as a criterion and Li et al. suggested that ICC greater than 0.75 indicates good to excellent reliability. However, it is stated that there are no standard values for acceptable reliability using ICC, as low ICC also reflects a lack of variability among the sampled subjects, the small number of subjects and raters. Therefore, given the current available publications and their limited number of sample size, we cannot come up with a reliable threshold value. We will mention this as a limitation in the “Impacts of patient populations and sample diversity on model accuracy comparison” part under the “Comparison on the study design” section.*

*Changes in the text:*

*We have added the corresponding statements in our manuscript. (Page 11, lines 336-344)*

*Comment 6:*

*would be useful to give some insights of the average human sensitivity/specificity*

*Reply 6:*

*Thank you for your comment. From our paper, line 236, in the study by Graby et al., the AI model identified 27% of undiagnosed aortic dilatation in baseline clinical reports or electronic patient records, achieving a sensitivity of 77% and a specificity of 99%. (23). This indicates that the AI model could assist human readers in diagnosing previously missed cases. Fink et al. found that their model can achieve 88.1% sensitivity and 95.7% specificity in thoracic aortic dissection*

*detection (20) (please refer to table 3). Therefore, we think that the current AI models can already achieve reasonable sensitivity and specificity but it is possible to come up with threshold if the sample diversity and sizes are enough. We will mention this in the “Sensitivity and specificity in TAAD diagnosis” part under the “Comparison on performance metrics of AI models” section.*

Changes:

*We have added the corresponding statement in our manuscript. (Page 8, lines 246-249)*

Comment 7:

*emphasize the clinical application and impact of ai in this specific topic*

Reply 7:

*Thank you for your comment. In the conclusion, we have summarized the benefits of applying AI in clinical settings, and list out the possible ways to deal with the current limitations in clinical scenarios.*

Changes:

*We have added the corresponding statements in our manuscript. (Pages 15-16, lines 472-484)*

Comment 8:

*Instead of a general statement about future applications, mention specific clinical outcomes AI could impact, such as reducing diagnostic time and increasing early intervention rates*

Reply 8:

*Thank you for your comment. We have specified the beneficial clinical outcomes contributed by AI, such as improving reporting efficiency and reducing rates of undiagnosed or misdiagnosed cases.*

Changes:

*We have added the corresponding statement in our manuscript. (Pages 15-16, lines 472-486)*