Validation of an AI-based algorithm for measurement of the thoracic aortic diameter in low-dose chest CT

Hamelink, I; de Heide, E; Pelgrim, G J; Kwee, T C; van Ooijen, P M A; de Bock, G H; Vliegenthart, R

Published in:
European Journal of Radiology

DOI:
10.1016/j.ejrad.2023.111067

European Journal of Radiology 167 (2023) 111067

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Validation of an AI-based algorithm for measurement of the thoracic aortic diameter in low-dose chest CT

I. (Iris) Hamelink [a], E. (Erik Jan) de Heide [a], G.J. (Gert Jan) Pelgrim [a], T.C. (Thomas) Kwee [a], P.M.A. (Peter) van Ooijen [b,c], G.H. (Truuske) de Bock [d], R. (Rozemarijn) Vliegenthart [a,*]

[a] Department of Radiology, University of Groningen, University Medical Center of Groningen, 9713GZ Groningen, The Netherlands
[b] Department of Radiation Oncology, University of Groningen, University Medical Center of Groningen, 9713GZ Groningen, The Netherlands
[c] Data Science in Health (DASH), University of Groningen, University Medical Center of Groningen, 9713GZ Groningen, The Netherlands
[d] Department of Epidemiology, University of Groningen, University Medical Center of Groningen, 9713GZ Groningen, The Netherlands

ARTICLE INFO

Keywords:
Thoracic aortic aneurysm
Chest CT
Artificial intelligence

ABSTRACT

Objectives: To evaluate the performance of artificial intelligence (AI) software for automatic thoracic aortic diameter assessment in a heterogeneous cohort with low-dose, non-contrast chest computed tomography (CT).
Materials and methods: Participants of the Imaging in Lifelines (ImaLife) study who underwent low-dose, non-contrast chest CT (August 2017–May 2022) were included using random samples of 80 participants <50y, ≥80y, and with thoracic aortic diameter ≥40 mm. AI-based aortic diameters at eight guideline compliant positions were compared with manual measurements. In 90 examinations (30 per group) diameters were reassessed for intra- and inter-reader variability, which was compared to discrepancy of the AI system using Bland-Altman analysis, paired samples t-testing and linear mixed models.
Results: We analyzed 240 participants (63 ± 16 years; 50 % men). AI evaluation failed in 11 cases due to incorrect segmentation (4.6 %), leaving 229 cases for analysis. No difference was found in aortic diameter between manual and automatic measurements (32.7 ± 6.4 mm vs 32.7 ± 6.0 mm, $p = 0.70$). Bland-Altman analysis yielded no systematic bias and a repeatability coefficient of 4.0 mm for AI. Mean discrepancy of AI (1.3 ± 1.6 mm) was comparable to inter-reader variability (1.4 ± 1.4 mm); only at the proximal aortic arch showed AI higher discrepancy (2.0 ± 1.8 mm vs 0.9 ± 0.9 mm, $p < 0.001$). No difference between AI discrepancy and inter-reader variability was found for any subgroup (all: $p > 0.05$).
Conclusion: The AI software can accurately measure thoracic aortic diameters, with discrepancy to a human reader similar to inter-reader variability in a range from normal to dilated aortas.

1. Introduction

Thoracic aortic aneurysms are usually detected incidentally and generally grow without symptoms until they rupture, which is often fatal [1]. Increased utilization of computed tomography (CT) results in detection of aortic aneurysms on scans that have been performed for other indications [2], e.g. lung cancer screening [3,4]. Aortic diameters as measured on chest CT serve as a reliable tool for diagnosis of thoracic aortic aneurysms, and is recommended by several guidelines (European Society of Cardiology (ESC) and American Heart Association (AHA)) [5,6]. However, guideline compliant analysis [7] of the thoracic aorta is a time consuming task, since it entails manual measurement of maximum and orthogonal diameters perpendicular to the lumen at nine positions. Potentially, this task could be carried out by artificial intelligence (AI) software integrated in the chest CT analysis workflow. Previous work [8] evaluated the performance of a commercial AI algorithm (AI-Rad companion) on 250 chest CT and CT angiography (CTA) scans from a diverse clinical cohort. They showed that this AI software could accurately measure thoracic diameters compared to a human reader, but a high coefficient of repeatability of 8.0 mm implied that human supervision remains needed. To determine if AI software for aortic evaluation could entirely replace human reading, the performance of AI software should be compared to the current clinical practice and particularly to the intra- and inter-reader variability present in

* Corresponding author.
E-mail addresses: i.hamelink@umcg.nl (I.(I. Hamelink), e.de.heide@umcg.nl (E.(E.J. de Heide), g.j.pelgrim@umcg.nl (G.J.(G.J. Pelgrim), t.c.kwee@umcg.nl (T.C.(T. Kwee), p.m.a.van.ooijen@umcg.nl (P.M.A.(P. van Ooijen), g.h.de.bock@umcg.nl (G.H.(T. de Bock), r.vliegenthart@umcg.nl (R.(R. Vliegenthart).

https://doi.org/10.1016/j.ejrad.2023.111067
Received 29 July 2023; Received in revised form 24 August 2023; Accepted 25 August 2023
Available online 26 August 2023

routine clinical setting. If the performance of the AI software is similar to manual aortic measurements and variability between human readers, this would indicate sufficient reliability for clinical implementation. AI software may have challenges on non-triggered, low-dose chest CT scans, as used in lung cancer screening [9], which necessitates separate validation. Furthermore, AI software needs to have sufficient accuracy across the range of aortic measurements that could be found in a general cohort, including those with a dilated aorta.

Therefore, the aim of this study was to evaluate the performance of AI software for automatic diameter assessment of the thoracic aorta at low-dose non-contrast chest CT, compared to manual evaluation, in a cohort with a range of aortic diameters.

2. Materials and methods

2.1. Study population and CT acquisition

For this study, a subset of participants from the Imaging in Lifelines (ImaLife) study was selected. The ImaLife study is embedded in the Lifelines cohort. Lifelines is a multi-disciplinary prospective population-based cohort study examining health and health-related behaviors of 167,729 persons living in the north of the Netherlands, established in 2006 [10]. It employs a broad range of investigative procedures in assessing the factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. From the Lifelines cohort, 12,128 participants aged 45 years and above were included for the ImaLife substudy, after attending the second Lifelines examination round (2014–2017) [11]. The ImaLife study was approved by the medical ethics committee of the University Medical Center Groningen and all participants provided informed consent. The ImaLife study was registered with the Dutch Central Committee on Research Involving Human Subjects (https://www.toetsingonline.nl, NL58592.042.16).

ImaLife participants underwent a non-triggered, non-contrast low-dose chest CT on a third-generation dual-source CT system (SOMATOM Force, Siemens Healthineers (Germany)) between August 2017 and May 2022. The scanning protocol included the following parameters: 120 kVp, 20 mAs, pitch 3.0, 1/0.7 mm slice thickness, and soft tissue kernel (Br40) reconstruction. Further parameters for CT acquisition and reconstruction were described in the design paper [11].

A purposive sample based on age and thoracic aortic diameter was made to ensure the full range of aortic diameters and age-related variations was present for this substudy: (A) 80 participants <50 years of age (low probability for calcifications, elongation and dilatation), (B) 80 participants ≥80 years of age (high probability for calcifications, elongation and dilatation), and (C) 80 participants with thoracic aortic diameter ≥40 mm as incidentally detected. Participants were first selected based on aortic diameter ≥40 mm, thereafter age, and finally sex (50 % of each).

Participant characteristics including cardiovascular comorbidities and risk factors were obtained from the Lifelines study and included history of coronary interventions (percutaneous coronary intervention or bypass surgery), congestive heart failure and stroke, diabetes mellitus (fasting blood glucose >7.0 mmol/L or use of antidiabetic medication), hypertension (systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg or use of antihypertensive medication) and hypercholesterolemia (low-density lipoprotein cholesterol ≥3.5 mmol/L or use of lipid lowering medication). Information on smoking included pack years and current smoking status.

2.2. Manual diameter measurements

Chest CT scans were reviewed using *syngo*.via software (VB60, Siemens Healthineers (Germany)) for measurement of thoracic aortic diameters at eight anatomic landmark positions following AHA and ESC guidelines. Following AHA guidelines, external aortic diameters were reported. The ninth measurement position at the location of the abdominal aorta was excluded since many CTs did not include the upper abdomen and/or contained much noise. The eight evaluated positions are depicted in Fig. 2 and comprise [7]:

1) Sinus of Valsalva
2) Sinotubular junction
3) Mid ascending aorta
4) Proximal aortic arch
5) Mid aortic arch
6) Proximal descending aorta
7) Mid descending aorta
8) Aorta at diaphragm

A licensed technical physician (IH) was trained by examining 60 cases under supervision of an EBCR radiologist (RV) with 15+ years of experience. The trained reader measured diameters at all landmark positions on the 240 chest CTs. A subset of 90 scans (30 per subset group) was re-evaluated in random order by the same reader after a period of one month to obtain intra-reader variability. This subset was also reviewed by a senior radiology resident (EH) (4 years of experience) to obtain inter-reader variability. Both reviewers were blinded to AI-based thoracic diameters.

2.3. Automatic diameter measurements by AI

The chest CT scans were also reviewed by AI software, the AI-Rad Companion chest CT research application (Version 08/2022, Siemens Healthineers (Germany)). Within the algorithm, aortic landmark positions were detected based on Deep Reinforcement learning [12] and segmentation of the aorta was done using an adversarial DI2IN in a symmetric convolutional encoder-decoder architecture [13]. It subsequently determined the appropriate locations to compute the diameters in the planes perpendicular to the centerline [14] and automatically generated those diameters. The model was trained on over 1000 non-cardiac chest CT examinations from multiple scanner types for suspected noncardiac pathology, before releasing the application to the market. The technical physician visually reviewed all assessments by the AI-Rad Companion for accuracy of segmentation, measurement position within the thoracic aorta and plane angulation. This step only included visual scoring, no adjustments or corrections were made to AI measurements.

2.4. Statistical analysis

Continuous variables were represented as mean ± standard deviation (SD) for overall diameter measurements, measurements per landmark position and per subgroup. Only maximum diameter measurements per landmark position per individual participant were analyzed, due to their clinical relevance. We visualized data to identify potential outliers. Intra- and inter-reader variability were derived from manual diameter measurements and represented as absolute measurement difference. Discrepancy between automatic and manual diameters was defined as the mean absolute difference ± SD.

AI-based diameters were compared to manual diameters using a linear mixed model with case as random effect and a fixed effect consisting of location ($p < 0.05$). Additionally, reproducibility and systematic error between automatic and manual measurements was assessed using Bland-Altman analysis.

Since we defined the maximum acceptable clinical bias of automated measurements to be equal to variability present in manual measurements, further analysis included comparison of AI-based absolute discrepancy and inter-reader variability. Comparison per landmark position included paired samples t-tests ($p < 0.00625$) and comparison per subgroup was done using linear mixed model with case as random effect and location as fixed effect ($p < 0.05$), both with Bonferroni correction.

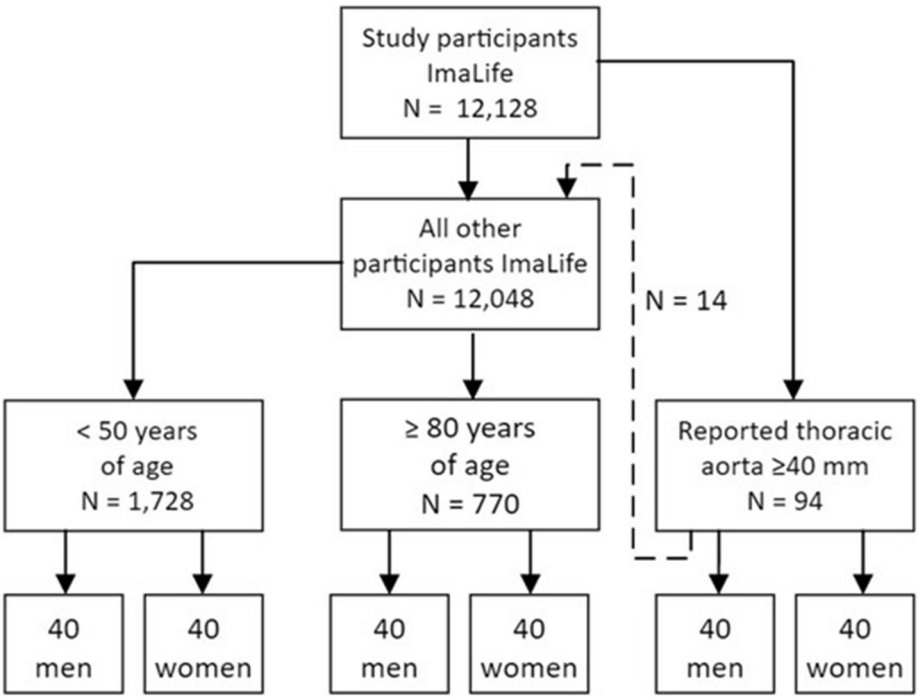

Fig. 1. Flow diagram of participant and subgroup selection from the ImaLife cohort.

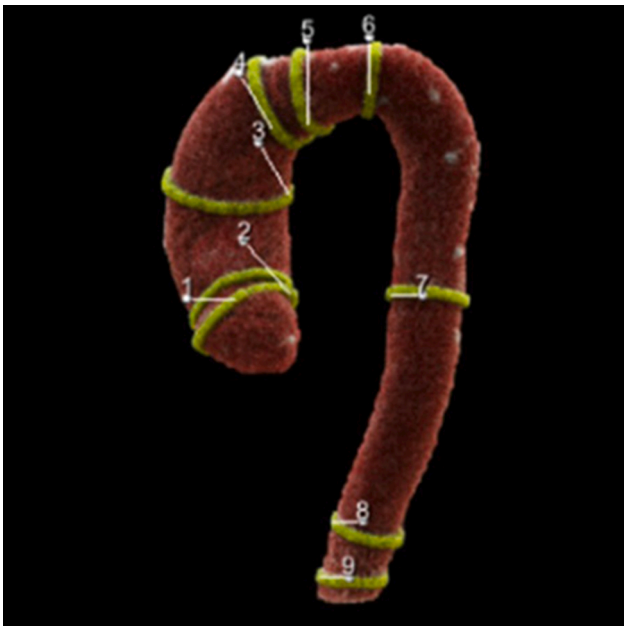

Fig. 2. Output from AI software showing the thoracic aorta with the eight guideline compliant diameter measurements (1: Sinus of Valsalva; 2: Sino-tubular junction; 3: Mid ascending aorta; 4: Proximal aortic arch; 5: Mid aortic arch; 6: Proximal descending aorta; 7: Mid descending aorta; 8: Aorta at diaphragm).

Additionally, the performance of the AI-based algorithm for the different subgroups was evaluated using the linear mixed model.

All statistical analyses were performed using SPSS version 28.0.0 (SPSS, Inc., Chicago, IL).

3. Results

3.1. Participant characteristics

A total of 240 participants (63 ± 16 years; 50 % men) were included; the selection process is depicted in Fig. 1. Demographic data including risk factors and comorbidities are shown in Table 1. Manual assessment of the thoracic aorta resulted in 1920 individual measurements for maximum diameters. Overall, the mean maximum aortic diameter was 32.7 ± 6.4 mm. Table 2 shows the mean maximum diameter per subgroup for every landmark position.

3.2. Manual diameter measurements

Overall absolute intra- and inter-reader variability was 0.8 ± 0.8 mm and 1.4 ± 1.4 mm, respectively. Fig. 3a and 3b show intra- and inter-reader variability at the eight aortic positions and per cohort subgroup. The systematic bias for the intra- and inter-reader analysis was

Table 1
Demographics data of the study population.

Parameter	<50 years of age	≥80 years of age	Thoracic aortic diameter≥40 mm
Number of participants	80	80	80
Age (years) ± SD*	46 ± 1	82 ± 2	59 ± 16
Male	40 (50 %)	40 (50 %)	40 (50 %)
Body mass index ± SD	25.7 ± 3.7	25.8 ± 2.8	26.3 ± 3.3
Cardiovascular comorbidities and risk factors			
Ever smoking	41 (51.2 %)	46 (57.5 %)	40 (50 %)
Current smoking	11 (13.8 %)	4 (5 %)	4 (5 %)
Pack years ± SD	13.8 ± 13.7	10.5 ± 9.7	12.4 ± 11.5
Coronary intervention	0 (0 %)	2 (2.5 %)	2 (2.5 %)
Congestive heart failure	0 (0 %)	1 (1.3 %)	0 (0 %)
Cerebrovascular disease	0 (0 %)	1 (1.3 %)	1 (1.3 %)
Diabetes mellitus	1 (1.3 %)	3 (3.8 %)	1 (1.3 %)
Hypertension	14 (17.5 %)	43 (53.8 %)	29 (36.3 %)
Hypercholesterolemia	19 (23.8 %)	53 (66.3 %)	31 (38.8 %)

*SD = standard deviation.

Table 2
Mean manual diameter measurements of the study population.

Diameter measurements* (mean ± SD**)	<50 years of age	≥80 years of age	Thoracic aortic diameter≥40 mm
Overall	29.3 ± 5.2	33.6 ± 5.1	35.0 ± 7.3
Sinus of Valsalva	36.0 ± 4.2	38.3 ± 3.8	40.7 ± 4.7
Sinotubular junction	31.3 ± 3.2	33.9 ± 3.4	37.6 ± 4.1
Mid ascending	33.5 ± 4.0	38.8 ± 3.9	44.2 ± 5.0
Proximal aortic arch	31.9 ± 3.0	37.0 ± 3.4	39.0 ± 3.6
Mid aortic arch	28.2 ± 2.4	32.4 ± 2.5	32.6 ± 3.1
Proximal thoracic descending	26.2 ± 2.3	31.5 ± 3.1	31.2 ± 5.1
Mid thoracic descending	23.9 ± 2.1	29.1 ± 3.4	28.0 ± 4.9
Thoracic aorta at diaphragm	23.6 ± 2.0	27.8 ± 2.7	27.1 ± 3.8

*Based on manual evaluation by reader 1.
**SD = standard deviation.

−0.2 mm and 0.7 mm respectively, with a repeatability coefficient of 2.2 mm and 3.5 mm respectively, see Supplemental Fig. 1.

3.3. Automatic diameter measurement by AI versus human reader

AI analysis failed in 11/240 cases (4.6 %) due to incomplete segmentation, and was complete in 229 participants, resulting in 1832 maximum diameter measurements. Visual check of failed cases showed 2 participants with aortic elongation. Dose Length Product and mean Heart Rate during scanning gave no indication for AI analysis failing. Mean BMI was slightly higher for the failed cases when compared to cases in which the automated software was able to produce results (26.4 ± 4.9 vs 25.9 ± 3.2). In outlier analysis one AI measurement of 130 mm at the proximal descending aorta was discarded due to incorrect plane angulation.

Mean maximum AI-based thoracic aortic diameter was 32.7 ± 6.0 mm versus 32.7 ± 6.4 mm for manual measurement ($p = 0.70$), and mean absolute discrepancy was 1.3 ± 1.6 mm. Bland-Altman analysis of AI measurements compared to manual measurements resulted in 0 mm systematic bias with a coefficient of repeatability of 4.0 mm (Fig. 4). For comparison, the systematic bias for the inter-reader analysis (in the subsample) was 0.7 mm with a repeatability coefficient of 3.5 mm.

3.4. Guideline compliant position analysis

Fig. 3a shows the mean absolute discrepancy for maximum aortic diameters between AI and manual measurement. In 50 % of the landmark positions, the mean absolute discrepancy was around 1 mm. The discrepancy of the AI system (2.0 ± 1.8 mm) exceeded the inter-reader variability (0.9 ± 0.9 mm) at the position of the proximal aortic arch ($p < 0.001$). In Fig. 5, an example of diameter measurements of reader 1, reader 2 and AI at that position is shown. Conversely, discrepancy of AI measurement to the human reader at the proximal descending aorta and the diaphragm position was significantly smaller than the inter-reader variability at those locations ($P < 0.001$); an example is shown in Fig. 6. The other aortic positions showed no difference when compared to the inter-reader variability.

3.5. Subgroup analysis

Mean AI-based and manual diameters were found to be respectively 29.3 ± 5.2 mm and 29.6 ± 4.9 mm for the young age group, 33.6 ± 5.1 mm and 33.5 ± 4.8 mm for the old age group, and 35.0 ± 7.3 mm and 35.0 ± 6.7 mm for the group with a thoracic aortic diameter ≥40 mm (all: $p > 0.05$ for comparing AI to manual measurement). Fig. 3b shows the discrepancies between the AI software and manual measurements alongside the inter-reader variability. AI-based measurements in participants with known thoracic aortic diameter ≥40 mm (1.6 ± 2.0 mm) showed higher absolute discrepancy than measurements in the groups with participants <50 y (1.2 ± 1.3 mm) and >80 y (1.2 ± 1.3 mm) (both $p < 0.001$).

4. Discussion

In 240 participants of a population-based study we showed that AI-based automated measurements of thoracic aortic diameters on low-dose, non-contrast CT were similar to those by manual assessment. Mean discrepancy of AI to the human reader fell within 2 mm. In seven landmark positions discrepancy of AI was similar to or even smaller than inter-reader variability of two human readers.

Previous studies also evaluated the performance of this commercial AI software. In 371 ECG-gated CTA scans, Pradella et al. showed that automated AI measurements differed <5 mm from human reading in 79.6 % of measurements [15]. Additionally, Artzner et al. showed no significant difference between AI measurements and human readers for any of the landmark positions in a dataset consisting of 122 CT and CTA scans [16]. In a small clinical CTA dataset (N = 18), Rueckel et al. [17] demonstrated relatively high discrepancy at the sinus of Valsalva and the proximal aortic arch, similar to where we found the largest discrepancy. They also showed the potential benefit when applying the software to follow-up scan protocols. The evaluation of such exams requires consistent measurement of aortic diameters along the aorta, in particular concerning positioning and angulation of diameter measurements. In their study, the automated software was able to reduce evaluation time without reducing follow-up assessment quality. However, low-dose, non-contrast chest CT as used in lung cancer screening was not included in these evaluations thus hampering a direct performance comparison with our study.

Overall, discrepancy of AI versus trained human readers in aortic sizing was low. The discrepancy at the position of the proximal aortic arch was relatively high when compared to the inter-reader variability (2.0 vs 0.9 mm). This might be due to the fact that correct positioning of this measurement location is prone to interpretation, and differed in a more systematic manner between the human reader and the AI system. The guideline [6] states that measurement should be placed at the origin of the innominate artery. In practice, it was found that the human readers chose the position slightly more proximally than AI. However, mean absolute discrepancy of AI only exceeded mean inter-reader variability by 1 mm, up to 2 mm, and therefore its effect in clinical practice can be considered negligible.

We showed that AI performance remained consistent in the old age group compared to the young age group. This indicates that the results of the AI system were not significantly affected by the potential influences of ageing on the thoracic aorta. This includes calcifications, elongation and dilation of the thoracic aorta. Discrepancies for participants with a thoracic aorta ≥40 mm slightly exceeded those of the young and old age group, indicating a slightly lower reproducibility of results in that subgroup. However, the difference in mean discrepancy was smaller than 0.5 mm and therefore not deemed clinically relevant.

A strength of our study was direct comparison of AI discrepancy to the human reader with inter- and intra-reader variability. Therefore, potential bias in the AI system could be interpreted in the right context. The results for intra- and inter-reader variability within our study are in line with previous literature [18] showing an inter-reader variability of 1–2 mm for the first four landmark positions. Importantly, the AI-based discrepancy in our study generally did not exceed 2 mm either, and was smaller for many landmark positions. Another strength was the purposive sampling of our dataset, which ensured a broad range of potential diameters and variations in thoracic aortic anatomy. This mitigates the risk that the AI system may perform worse in a certain part of the population, and particularly those with diseased aortas. The main advantage of such a population-based series is that participants within this study were a sample of the potential target population of future lung cancer screening and therefore the results give an indication of AI performance if it would be implemented in such a setting.

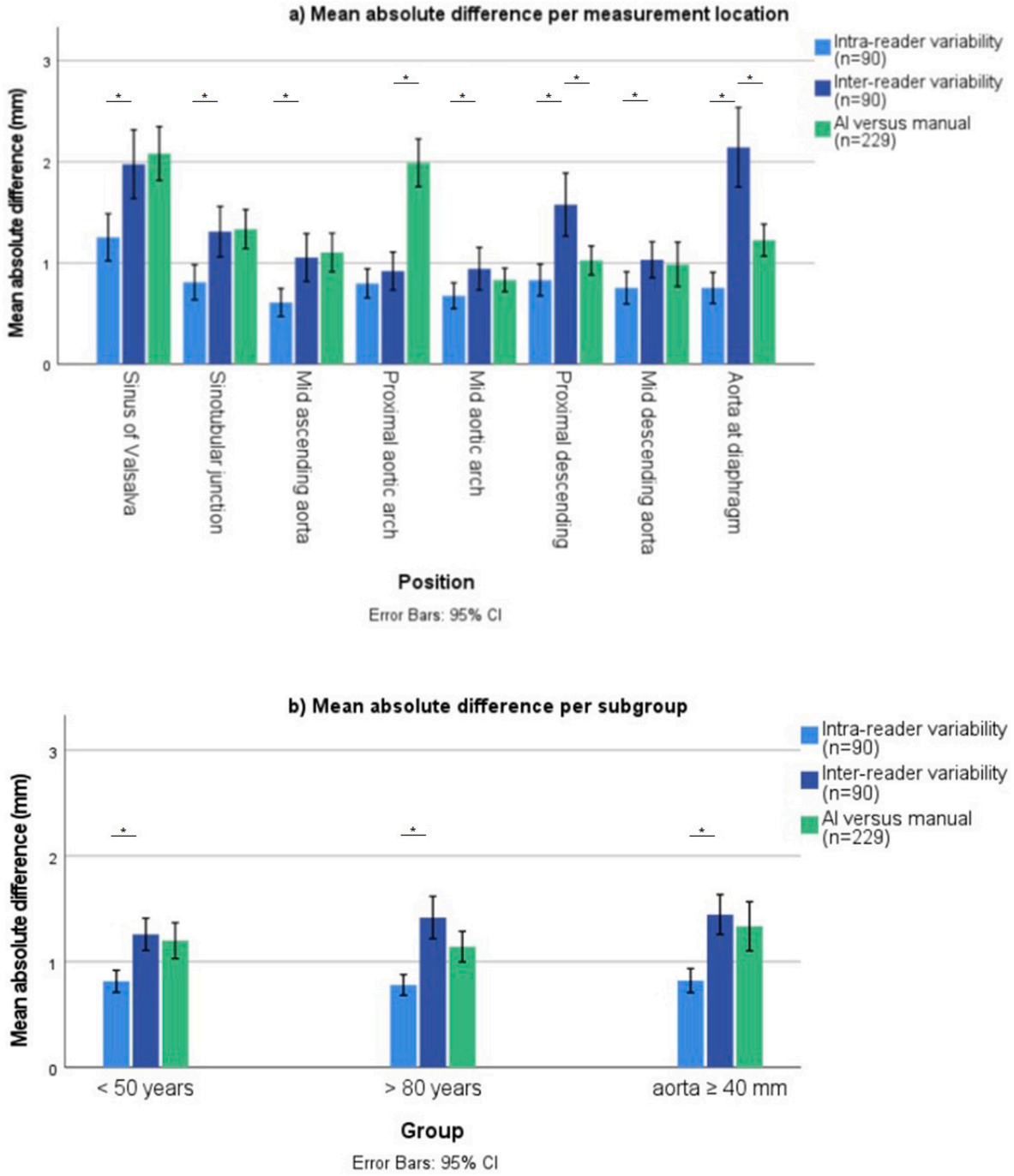

Fig. 3. Mean absolute difference for the landmark positions overall (Fig. 3a) and per subgroup (Fig. 3b). Blue bars represent intra- and inter-reader variability for the manual maximum diameter measurements in a subgroup of 90 CT scans. Green bars represent discrepancy between AI software measurements and manual measurements obtained from entire sample (229 CT scans) (* $p < 0.05$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A limitation of our study was the fact that the two trained readers in our study were not (cardiovascular) radiologists, similarly to previous work [8]. However, Yacoub et al. showed that good-to-excellent inter-reader agreement for thoracic aortic diameter evaluation between a non-experienced reader and experienced radiologists [18]. Secondly, the set-up of our CT data review system did not allow for comparison of time involvement. However, in the study by Rueckel et al. a time reduction of 63 % was found using this AI software [17]. In the context of our study and its results, these reporting times could potentially be even further reduced by using the AI-based algorithm as automatic source for pointing out conspicuous cases during lung cancer screening without the need and possibility to check (and measure) every scan.

In conclusion, in low-dose chest CT, AI software can accurately measure guideline compliant thoracic aortic diameters in a range from normal to dilated aortas, with discrepancy to a human reader similar to inter-reader variability. This suggests that AI can replace human measurement of aortic diameters in adequately segmented cases, for example in lung cancer screening.

Funding

The ImaLife study is supported by an institutional research grant

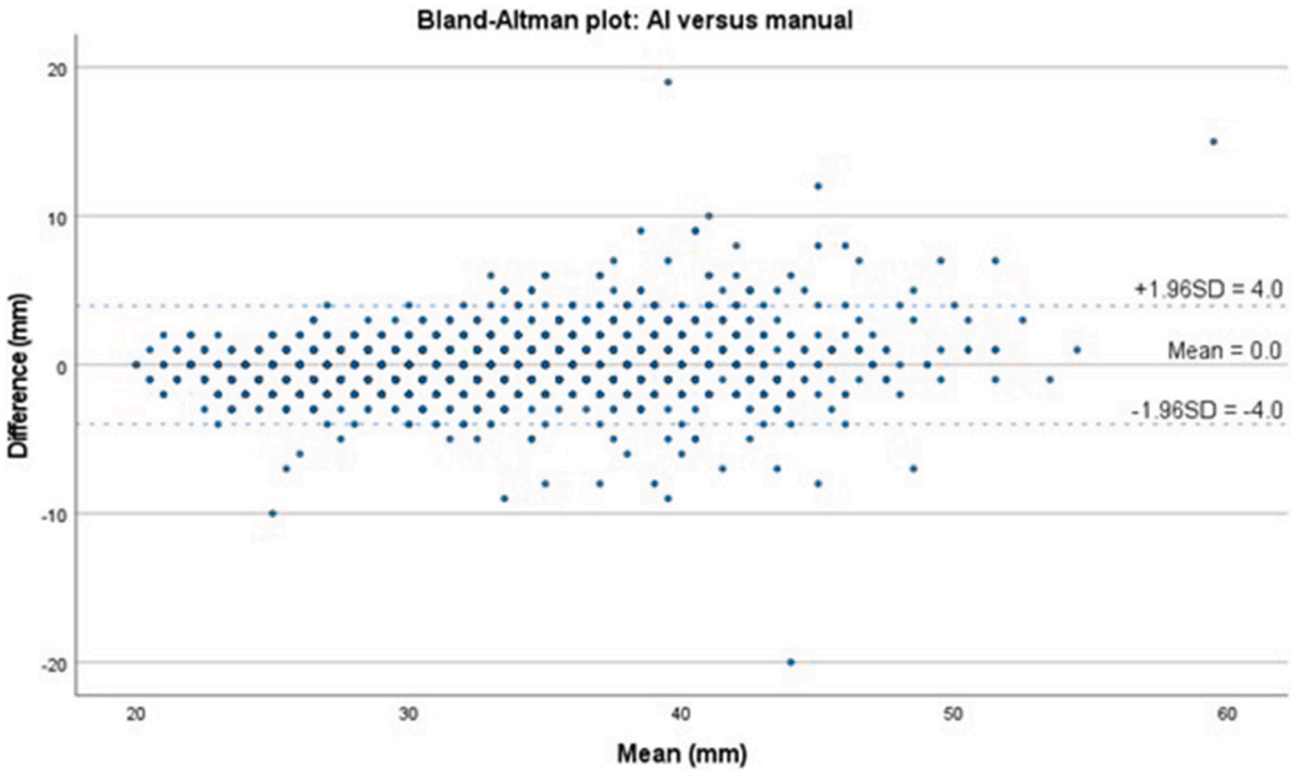

Fig. 4. Bland-Altman plot representing the agreement between AI and manual measurement in 229 participants.

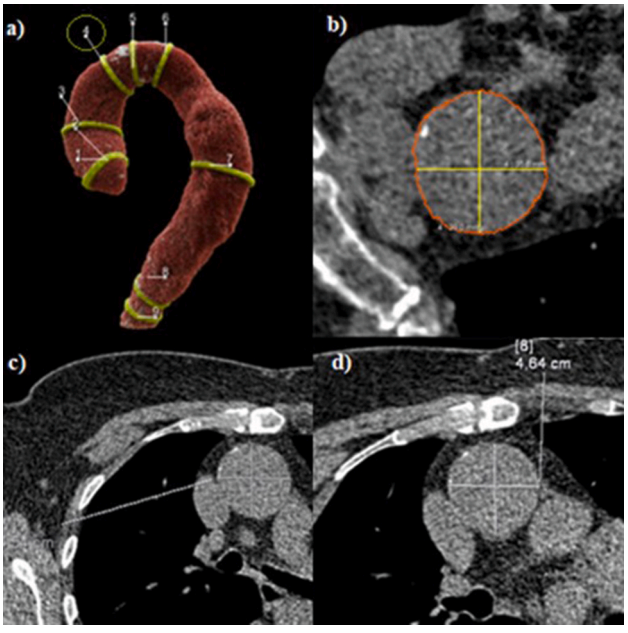

Fig. 5. Output of the AI software (a and b) next to manual measurement by reader 1 (c) and reader 2 (d). It represents a case (female, 81 y) where AI showed relatively large discrepancy (6 mm) at the proximal ascending aorta (nr. 4, circled) due to erroneous positioning compared to the inter-reader variability (1 mm).

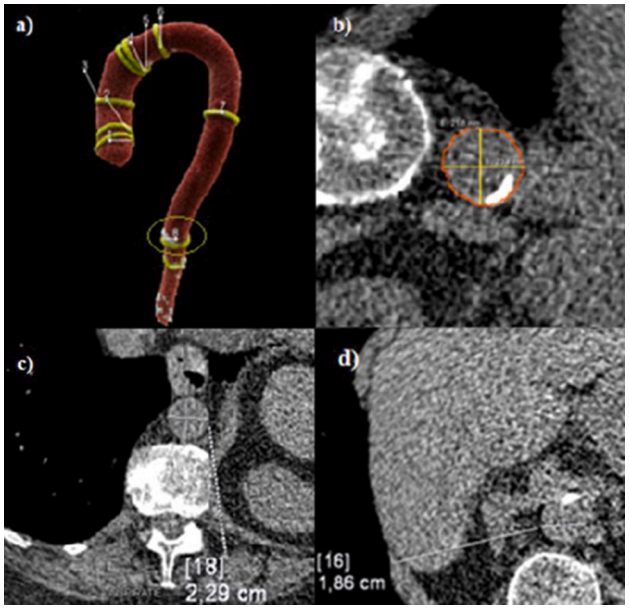

Fig. 6. Output of the AI software (a and b) next to manual measurement by reader 1 (c) and reader 2 (d). It represents a case (female, 80 y) where inter-reader variability at the aortic diaphragm (nr. 8, circled) was relatively high (4 mm), but AI only slightly differed from reader 1.

from Siemens Healthineers and by the Ministry of Economic Affairs and Climate Policy by means of the PPP Allowance made available by the Top Sector Life Sciences & Health to stimulate public–private partnerships. No funding body had any role in study design, data collection, analysis, interpretation, or writing of the manuscript, nor in the decision to publish.

CRediT authorship contribution statement

I. (Iris) Hamelink: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **E. (Erik Jan) de Heide:** . **G.J. (Gert Jan) Pelgrim:** . **T.C. (Thomas) Kwee:** . **P.M.A. (Peter) van Ooijen:** . **G.H. (Truuske) de Bock:** . **R. (Rozemarijn) Vliegenthart:** Writing – review & editing,

Supervision, Resources, Project administration, Methodology, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Rozemarijn Vliegenthart has received an institutional research grant from Siemens Healthineers. The PhD project of Iris Hamelink is supported by a PUSH grant from Siemens Healthineers. The remaining authors have no disclosures to report.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrad.2023.111067.

References

[1] E. Gouveia, R. Melo, G. Silva Duarte, A. Lopes, M. Alves, D. Caldeira, R. Fernandes E Fernandes, L. Mendes Pedro, Incidence and prevalence of thoracic aortic aneurysms: a systematic review and meta-analysis of population-based studies, Semin. Thorac. Cardiovasc. Surg. 34 (1) (2022) 1–16.

[2] G.A. Kuzmik, A.X. Sang, J.A. Elefteriades, Natural history of thoracic aortic aneurysms, J. Vasc. Surg. 56 (2) (2012) 565–571.

[3] National Lung Screening Trial Research Team, D.R. Aberle, A.M. Adams, C.D. Berg, W.C. Black, J.D. Clapp, R.M. Fagerstrom, I.F. Gareen, C. Gatsonis, P.M. Marcus, J. D. Sicks, Reduced lung-cancer mortality with low-dose computed tomographic screening, New Engl. J. Med. 365 (5) (2011) 395–409.

[4] H.J. de Koning, C.M. van der Aalst, P.A. de Jong, E.T. Scholten, K. Nackaerts, M. A. Heuvelmans, J.J. Lammers, C. Weenink, U. Yousaf-Khan, N. Horeweg, S. van 't Westeinde, M. Prokop, W.P. Mali, F.A.A. Mohamed Hoesein, P.M.A. van Ooijen, J. G.J.V. Aerts, M.A. den Bakker, E. Thunnissen, J. Verschakelen, R. Vliegenthart, J. E. Walter, K. Ten Haaf, H.J.M. Groen, M. Oudkerk, Reduced lung-cancer mortality with volume CT screening in a randomized trial, N. Engl. J. Med. 382 (6) (2020) 503–513.

[5] R. Erbel, V. Aboyans, C. Boileau, E. Bossone, R.D. Bartolomeo, H. Eggebrecht, A. Evangelista, V. Falk, H. Frank, O. Gaemperli, M. Grabenwöger, A. Haverich, B. Lung, A.J. Manolis, F. Meijboom, C.A. Nienaber, M. Roffi, H. Rousseau, U. Sechtem, P.A. Sirnes, R.S. Allmen, C.J. Vrints, ESC Committee for Practice Guidelines, 2014 ESC Guidelines on the diagnosis and treatment of aortic diseases: Document covering acute and chronic aortic diseases of the thoracic and abdominal aorta of the adult. The Task Force for the Diagnosis and Treatment of Aortic Diseases of the European Society of Cardiology (ESC), Eur. Heart J. 35 (41) (2014) 2873–2926.

[6] L.F. Hiratzka, G.L. Bakris, J.A. Beckman, R.M. Bersin, V.F. Carr, D.E. Casey Jr, K. A. Eagle, L.K. Hermann, E.M. Isselbacher, E.A. Kazerooni, N.T. Kouchoukos, B. W. Lytle, D.M. Milewicz, D.L. Reich, S. Sen, J.A. Shinn, L.G. Svensson, D. M. Williams, American College of Cardiology Foundation; American Heart Association Task Force on Practice Guidelines; American Association for Thoracic Surgery; American College of Radiology; American Stroke Association; Society of Cardiovascular Anesthesiologists; Society for Cardiovascular Angiography and Interventions; Society of Interventional Radiology; Society of Thoracic Surgeons; Society for Vascular Medicine, 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with thoracic aortic disease: a report of the American college of Cardiology Foundation/American Heart Association Task Force on practice guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society of Vascular Medicine, Circulation 121 (13) (2010) e266–e369 [Published correction appears in Circulation, 122(4): e410].

[7] M.J. van Hout, A.J. Scholte, J.F. Juffermans, J.J. Westenberg, L. Zhong, X. Zhou, S. M. Schalla, M.D. Hope, J. Bremerich, C.M. Kramer, M. Dewey, K.G. Ordovas, D. A. Bluemke, H.J. Lamb, How to measure the aorta using MRI: a practical guide, J. Magn. Reson. Imaging 52 (4) (2020) 971.

[8] C.B. Monti, M. van Assen, A.E. Stillman, S.J. Lee, P. Hoelzer, G.S.K. Fung, F. Secchi, F. Sardanelli, C.N. De Cecco, Evaluating the Performance of a Convolutional Neural Network Algorithm for Measuring Thoracic Aortic Diameters in a Heterogeneous Population. Radiology, Artif. Intell. 4 (2) (2022), e210196.

[9] J.M. Noothout, B.D. de Vos, J.M. Wolterink, I. Išgum, Automatic segmentation of thoracic aorta segments in low-dose chest CT, Med. Imaging: Image Process. 10574 (2018) 446–451.

[10] A. Sijtsma, J. Rienks, P. van der Harst, G. Navis, J.G. Rosmalen, A. Dotinga, Cohort Profile Update: Lifelines, a three-generation cohort study and biobank, Inter. J. Epidemiol. 51 (5) (2022) e295–e302.

[11] C. Xia, M. Rook, G.J. Pelgrim, G. Sidorenkov, H.J. Wisselink, J.N. van Bolhuis, P.M. A. van Ooijen, J. Guo, M. Oudkerk, H. Groen, M. van den Berge, P. van der Harst, H. Dijkstra, M. Vonder, M.A. Heuvelmans, M.D. Dorrius, P.P. De Deyn, G.H. de Bock, A. Dotinga, R. Vliegenthart, Early imaging biomarkers of lung cancer, COPD and coronary artery disease in the general population: rationale and design of the ImaLife (Imaging in Lifelines) Study, European J. Epidemiol. 35 (2020) 75–86.

[12] F.C. Ghesu, B. Georgescu, Y. Zheng, S. Grbic, A. Maier, J. Hornegger, D. Comaniciu, Multi-scale deep reinforcement learning for real-time 3D-landmark detection in CT scans, IEEE Trans. Pattern Anal. Mach. Intell. 41 (1) (2019) 176–189.

[13] A.M. Fischer, B. Yacoub, R.H. Savage, J.D. Martinez, J.L. Wichmann, P. Sahbaee, S. Grbic, A. Varga-Szemes, U. Schoepf, Machine learning/deep neuronal network: routine application in chest computed tomography and workflow considerations, J. Thorac. Imaging 35 (2020) S21–S27.

[14] M. Remy-Jardin, J.B. Faivre, R. Kaergel, A. Hutt, P. Felloni, S. Khung, A.L. Lejeune, J. Giordano, J. Remy, Machine learning and deep neural network applications in the thorax: pulmonary embolism, chronic thromboembolic pulmonary hypertension, aorta, and chronic obstructive pulmonary disease, J. Thorac. Imaging 35 (2020) S40–S48.

[15] M. Pradella, T. Weikert, J.I. Sperl, et al., Fully automated guideline-compliant diameter measurements of the thoracic aorta on ECG-gated CT angiography using deep learning, Quant. Imaging Med. Surg. 11 (10) (2021) 4245–4257.

[16] M. Pradella, T. Weikert, J.I. Sperl, R. Kärgel, J. Cyriac, R. Achermann, A.W. Sauter, J. Bremerich, B. Stieltjes, P. Brantner, G. Sommer, Assessing the accuracy of an artificial intelligence-based segmentation algorithm for the thoracic aorta in computed tomography applications, Diagnostics (Basel, Switzerland) 12 (8) (2022), 1790.

[17] J. Rueckel, P. Reidler, N. Fink, J. Sperl, T. Geyer, M.P. Fabritius, J. Ricke, M. Ingrisch, B.O. Sabel, Artificial intelligence assistance improves reporting efficiency of thoracic aortic aneurysm CT follow-up, Eur. J. Radiol. 134 (2021), 109424.

[18] B. Yacoub, R.E. Stroud, D. Piccini, U.J. Schoepf, J. Heerfordt, J. Yerly, L. Di Sopra, J.D. Rollins, D.A. Turner, T. Emrich, F. Xiong, P. Suranyi, A. Varga-Szemes, Measurement accuracy of prototype non contrast, compressed sensing-based, respiratory motion-resolved whole heart cardiovascular magnetic resonance angiography for the assessment of thoracic aortic dilatation: comparison with computed tomography angiography, J. Cardiovasc. Magn. Reson. 23 (1) (2021) 1–13.