

<https://doi.org/10.1038/s41746-024-01269-4>

Prospective clinical evaluation of deep learning for ultrasonographic screening of abdominal aortic aneurysms

Check for updates

I-Min Chiu^{1,2,3}, Tien-Yu Chen⁴, You-Cheng Zheng⁴, Xin-Hong Lin¹, Fu-Jen Cheng¹, David Ouyang^{2,5}✉ & Chi-Yung Cheng^{1,3,5}✉

Abdominal aortic aneurysm (AAA) often remains undetected until rupture due to limited access to diagnostic ultrasound. This trial evaluated a deep learning (DL) algorithm to guide AAA screening by novice nurses with no prior ultrasonography experience. Ten nurses performed 15 scans each on patients over 65, assisted by a DL object detection algorithm, and compared against physician-performed scans. Ultrasound scan quality, assessed by three blinded expert physicians, was the primary outcome. Among 184 patients, DL-guided novices achieved adequate scan quality in 87.5% of cases, comparable to the 91.3% by physicians ($p = 0.310$). The DL model predicted AAA with an AUC of 0.975, 100% sensitivity, and 97.8% specificity, with a mean absolute error of 2.8 mm in predicting aortic width compared to physicians. This study demonstrates that DL-guided POCUS has the potential to democratize AAA screening, offering performance comparable to experienced physicians and improving early detection.

Abdominal aortic aneurysm (AAA) is a life-threatening condition that contributes to a crude mortality rate of ~150,000–200,000 deaths per year worldwide¹. More than two-thirds of patients with a ruptured abdominal aortic aneurysm present without a prior diagnosis of AAA². An AAA is typically defined as aortic enlargement with a diameter of 3.0 cm or larger³. Early studies suggested that aortic diameters <3.0 cm did not necessitate regular follow-up. However, recent data reveal that 13.8% of patients with aortic diameters between 2.6 and 2.9 cm experienced growth to over 5.5 cm within 10 years^{4,5}. Early detection and treatment can considerably decrease AAA-related mortality, especially in the elderly⁶. Recent guidelines recommend ultrasonography for general AAA screening in at-risk populations, specifically men over 65, with a smoking history, with other peripheral aneurysms, a first-degree relative with an AAA, as well as those who have undergone organ transplantation⁷.

Ultrasonography, the most common imaging screening modality for AAA, has proven to be effective and less costly compared to standard computed tomography (CT) scans⁸. Nonetheless, ultrasound is generally performed by highly trained sonographers, and the interpretation is carried out by board-certified physicians. This challenge is further exacerbated by the restricted accessibility of sophisticated ultrasound equipment. Both of these factors contribute to ultrasound examinations having the longest

waiting times compared to other imaging modalities, thereby reducing the cost-effectiveness of general screening for AAA^{3–5}.

In the past decade, Point-of-Care Ultrasound (POCUS) has seen extensive use across various hospital settings, including outpatient clinics, emergency departments, wards, and operating rooms^{6,7}. It has played a crucial role in medically underserved areas, from rural regions in developed countries to low-income nations, and is even used in manned space flights⁸. Portable ultrasound machines have facilitated the acquisition of high-quality images suitable for diagnostic purposes⁹. This has the potential to enhance diagnostic capabilities, especially in remote areas where experienced sonographers may not be readily available¹⁰.

AI revolutionizes healthcare by rapidly interpreting images, detecting abnormalities, segmenting organs and lesions, and facilitating early disease identification¹¹. A prior study suggests that AI empowers individuals with no previous experience in ultrasonography to perform diagnostic trans-thoracic echocardiographic studies, encompassing the evaluation of left and right ventricular function, as well as the identification of pericardial effusion^{12–14}. In addition, AI can guide novices to capture satisfactory diagnostic images of the Morrison pouch during trauma examination^{15,16}. The objective of this study is to develop and validate a deep learning (DL) model that guides abdominal aorta scanning and to investigate its potential

¹Department of Emergency Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan. ²Department of Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA. ³Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan. ⁴Division of Cardiology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan. ⁵These authors contributed equally: David Ouyang, Chi-Yung Cheng. ✉e-mail: david.ouyang@cshs.org; qzsecawsxd@cgmh.org.tw

in assisting novices to obtain qualified scans. This approach could aid in the screening of potential AAA patients by advancing the detection of symptomatic aneurysms, screening asymptomatic AAAs in at-risk groups, and monitoring aneurysm growth until treatment is necessary.

Result

Patients at least 65 years old visiting the outpatient clinic of the Cardiology department at the studied hospital were recruited between June and August 2023. Ten registered nurses without prior experience performing or interpreting ultrasonography were recruited for the trial from hospital personnel. During the study period, 188 patients were included. Two patients were unable to lie flat, one patient was unable to provide informed consent, and one patient was unwilling to provide informed consent. A total of 184 patients completed both the nurse and physician examinations (Fig. 1). Their median (IQR) age was 72 (67–79), 57.1% were male, and the median (IQR) body mass index (BMI) was 25.1 (23.3–28.1). 131 (71.2%) of them had Hypertension, 64 (34.8%) had diabetes mellitus, 83 (45.1%) had heart disease, and 41 (22.3%) had a smoking history. AAA was diagnosed by physicians in 3 patients, representing 1.6% of the study cohort. Table 1 shows the demographics of the included patients.

Regarding the primary outcome, no significant difference was found in the rate of qualified videos between DL-guided scans 87.5% (95% CI: 82.7–92.3%), and physician-performed scans 91.3% (95% CI: 87.2–95.4%), with a *p*-value of 0.310. Furthermore, the qualified rate for DL guidance remained consistent across patients with varying BMI levels: 87.6% (95% CI:

80.9–94.1%) in patients with BMI > 25 and 87.4% (95% CI: 80.4–94.3%) in patients with BMI < 25. After five rounds of examination, the proficiency of scans slightly increased, reaching an 88.1% (95% CI: 82.3–93.9%) qualification rate, as detailed in Table 2.

Regarding primary outcome, no significant difference was found in the rate of qualified videos between DL-guided scans 87.5% (95% CI: 82.7–92.3%), and physician-performed scans 91.3% (95% CI: 87.2–95.4%), with *p*-value of 0.310. Furthermore, the qualified rate for DL guidance remained consistent across patients with varying BMI levels: 87.6% (95% CI: 80.9–94.1%) in patients with BMI > 25 and 87.4% (95% CI: 80.4–94.3%) in patients with BMI < 25. After 5 rounds of examination, the proficiency of scans slightly increased, reaching an 88.1% (95% CI: 82.3–93.9%) qualification rate, as detailed in Table 2.

The time to complete the study was longer with DL guidance, averaging 37 s (IQR 21–60), compared to 20 s (IQR 16–33) for physician-led scans (*p* < 0.001). For nurses using DL guidance, the completion time was longer in patients with a BMI over 25, taking 42 s (IQR 27–67), as opposed to 30 sec (IQR 18–54) for those with a BMI under 25. Physicians’ scans showed similar patterns between different levels of BMI. With the increased use of the DL system, nurses’ scan times decreased; after five scans, the average time was reduced from 53 sec (IQR 38–82) to 30.5 s (IQR 18–55), as reported in Table 2.

Of the 161 scans under DL guidance classified as diagnostic quality, the predicted maximal aortic widths showed a mean absolute error of 2.8 mm compared with physician measurements, as depicted in Fig. 2.

Fig. 1 | Study design. The study design involved nurse participation and patient enrollment. Ten nurses with no prior experience received ultrasound and deep learning guidance training and performed practice scans. A total of 184 patients were enrolled, scanned by both nurses and experts, with the deep learning model predicting maximal aortic width.

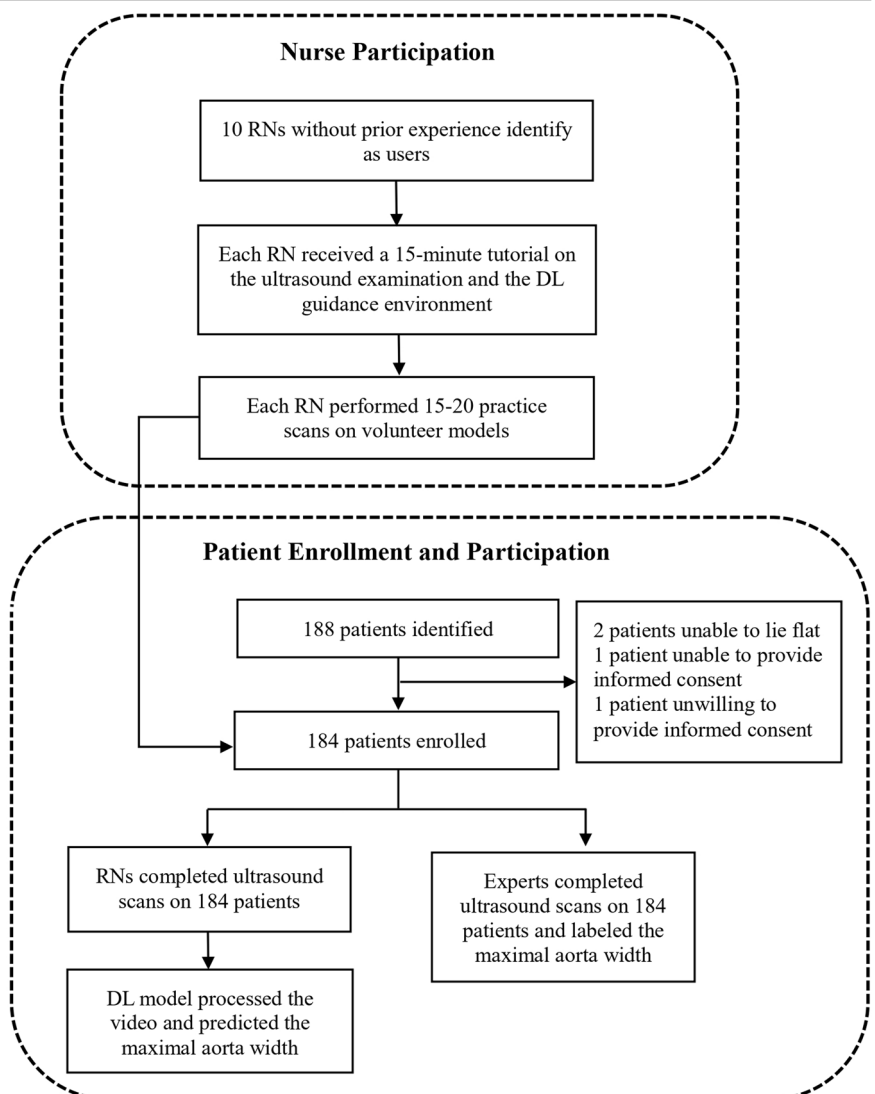


Table 1 | Demographic of included patients

	Median (IQR)/N (%)
Total patients	184
Age	72 (67–79)
Male	105 (57.1%)
BMI	25.1 (23.3–28.1)
<i>Underlying disease</i>	
Hypertension	131 (71.2%)
DM	64 (34.8%)
Heart disease	83 (45.1%)
PAOD	1 (0.5%)
CKD	24 (13.0%)
Abdominal operation	48(26.1%)
Smoking	41(22.3%)
<i>Family history</i>	
Heart disease	31 (16.8%)
Aortic disease	1 (0.5%)
AAA	3 (1.6%)

DM diabetes mellitus, PAOD peripheral artery occlusive disease, CKD chronic kidney disease.

Table 2 | Comparison of nurse-acquired and physician-acquired studies for primary and secondary clinical parameters

	Physician	DL guidance	p-value
Qualified video	No. (%) [95% CI]	No. (%) [95% CI]	
Total	168 (91.3) [87.2–95.4]	161 (87.5) [82.7–92.3]	0.310
BMI > 25 (N = 97)	91 (93.8) [87.6–99.8]	85 (87.6) [80.9–94.1]	0.137
BMI < 25 (N = 87)	77 (88.5) [81.5–95.4]	76 (87.4) [80.4–94.3]	0.816
Before 5 scans (N = 50)	45 (90.0) [81.7–98.3]	43 (86.0) [76.4–95.6]	0.613
After 5 scans (N = 134)	123 (91.8) [86.5–97.1]	118 (88.1) [82.3–93.9]	0.659
Time to complete study	Median (IQR) (s)	Median (IQR) (s)	
Total	20 (16–33)	37 (21–60)	<0.001
BMI > 25 (N = 97)	21 (16–35)	42 (27–67)	<0.001
BMI < 25 (N = 87)	18 (15–28)	30 (18–54)	<0.001
Before 5 scans (N = 50)	20 (16–35)	53 (38–82)	<0.001
After 5 scans (N = 134)	20 (15–30)	30.5 (18–55)	<0.001

Of these scans, 159 (98.8%) had a discrepancy of <1 cm when compared to physician labeling. 3 patients (1.6%) were diagnosed with AAA in this study based on the physician’s ultrasound findings. The DL model can predict AAA with an AUC of 0.975 (95% CI: 0.943–1). While using 2.5 cm as the cut-off threshold, the DL model identified 7 videos with suspected AAA, achieving 100% sensitivity in detecting AAA in these patients, along with 97.8% specificity, 42.8% positive predictive value, and 100% negative predictive value (Table 3). Manual inspection of the 4 false positive videos revealed that 2 videos were due to falsely enlarged bounding boxes, while the other 2 were due to the misprediction of non-aorta structures as the aorta, specifically a liver cyst and a fluid-containing bowel loop.

Discussion

In this study, we developed and validated a DL algorithm for guiding novice users performing AAA screening with POCUS, which notably

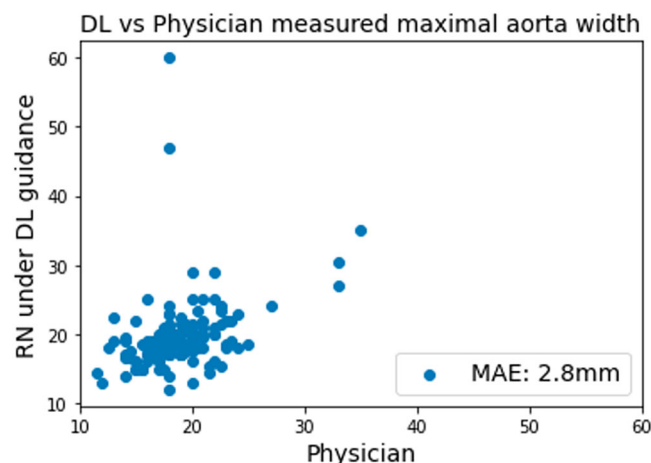


Fig. 2 | Scatter plot of maximal aortic width measurements by deep learning predictions and physician labels. Compared with physician measurements, the deep learning model demonstrated a mean absolute error of 2.8 mm.

Table 3 | Diagnostic performance for predicting AAA by DL algorithm

	DL-guidance
AUC	0.975 (0.943–1)
Sensitivity	100%
Specificity	97.8%
Positive predictive value	42.8%
Negative predictive value	100%

demonstrated parity with experienced physicians in producing diagnostic-quality videos. This finding suggests DL guidance can compensate for a lack of extensive sonography experience, potentially broadening AAA screening accessibility. Moreover, the DL model exhibited robust performance across a range of patient body mass indexes and showed a notable learning curve, with improved scan times after repeated use. The precision of our DL model in predicting the presence of AAA was quantitatively reflected by an AUC of 0.973 in our sample population. This high level of accuracy not only demonstrates the DL model’s capability to distinguish between normal and pathological findings but also suggests its potential as a reliable tool for early detection. These results underscore the practicality and efficiency of implementing AI in clinical ultrasound practice, which may help to reduce waiting time for ultrasound examination, especially in resource-limited settings where access to skilled sonographers is challenging.

The utility of POCUS is well recognized for its convenience and the immediacy with which it delivers diagnostic imaging at the patient’s bedside. However, its effective use is traditionally limited by the operator’s expertise, with the quality of the interpretation being heavily reliant on the sonographer’s experience. Current protocols require at least a 3-month training program for technicians to become familiar with examining the abdominal aorta using ultrasound¹⁷. This presents a significant challenge in resource-limited settings where access to highly trained professionals is scarce^{18,19}, which led to ongoing debate about the cost-effectiveness and relevance of AAA screening. Given its low prevalence, with historical data shows the prevalence rate of AAA 1.3–4.9% in selected risk populations, the balance between the costs and benefits of widespread screening is called into question^{20,21}. This is precisely where the integration of DL guidance in AAA screening could play a transformative role. By potentially reducing the time and expertise required for accurate screening, DL guidance can lower operational costs and improve the efficiency of

screening programs. Moreover, the enhanced accuracy of DL-guided screening might lead to more effective identification of AAA cases even in a landscape of lower prevalence, ensuring that resources are optimally utilized.

Advances in cardiovascular ultrasound interpretation using AI have been significant in recent years, with numerous studies demonstrating automated quantification of cardiac structures and function, and AI-driven disease identification showing less variability than semi-automated or manual analyses^{13,22–24}. The convergence of AI-guided acquisition with automated interpretation could expand ultrasound's reach, improving the recognition of pathology. Our study's DL algorithm addresses this by providing real-time guidance to novice users, effectively bridging the gap between the ease of use of POCUS and the expertise required for accurate diagnosis. By supplementing the user's limited experience with sophisticated AI, we facilitate a higher standard of care that could potentially revolutionize the screening process for conditions like AAA, where early detection is crucial.

Several potential explanations for our findings emerge upon examination of the DL algorithm's performance. The ability of the algorithm to offer real-time, continuous guidance likely played a professional guiding role in enabling novice operators to achieve a high rate of qualified videos. This real-time feedback is particularly useful during scanning, as it assists users in adjusting the ultrasound probe to optimize the visualization of the aorta. Such immediate guidance can improve image quality, a benefit observed in previous studies where AI support enhanced the outcomes of sonography procedures, such as in the Echocardiography¹⁴ or Focused Assessment With Sonography in Trauma exams¹⁶. The similarity between these findings and our own suggests a shared advantage of AI assistance across different ultrasound applications. The learning curve evidenced by the reduction in scan times with repeated use indicates that users are not only becoming more adept at operating the AI system but are also gaining a better understanding of the aortic structure. This suggests a synergistic enhancement in the operator's ability to acquire diagnostic images, pointing to a productive interaction between human users and the AI tool over time.

The architecture of our DL model, which delineates the abdominal aorta with a bounding box in each frame of the ultrasound video, serves two critical functions. First, it offers real-time feedback during the scanning process, guiding users in adjusting their probe to achieve optimal imaging. Second, it enables precise diagnostic measurements post-scan. By accurately capturing the anatomy of the aorta, our algorithm processes each frame to determine the aortic width, subsequently calculating the maximal width from the entire video. The minimal average error of 2.8 mm between the expert measurements and those obtained via DL guidance, along with the high AUC for detecting an AAA, attests to the DL model's effective training. For further evaluation, a threshold of 2.5 cm led to expert review for 9 (4.9%) of the 184 scans guided by DL. Within this group, three AAA cases were accurately identified, resulting in a sensitivity of 1.0 and a specificity of 0.94 in our cohort (Table 3). Notably, there were two scans where the discrepancy in maximal width measurements between expert interpretation and DL guidance exceeded 1 cm (Fig. 2). Manual review of these outliers showed that the DL algorithm had incorrectly identified a 6.0 cm liver cyst and a 4.8 cm fluid-filled small bowel loop. These cases of misclassification demonstrate that when videos with predictive diameters suspicious of AAA are scrutinized, physicians can readily discern true positives from false positives.

There are a few limitations of the study. First, the study was conducted in a controlled clinical setting in the outpatient clinic, which allowed access to a large patient population but may not fully replicate POCUS use conditions in remote or underserved areas. Second, while the sample size was sufficient for the primary endpoints, the relatively small number of patients and nurses limits the assessment of generalizability. Larger studies, particularly with more AAA cases, would yield more robust data. However, this study demonstrates proof-of-concept, showing the feasibility of AI guidance in performing abdominal aorta ultrasound scanning. Third, there was no control group for the nurse scanners. The comparison in the study was

against physician acquisitions, but an additional control group of novices untrained with the algorithm was not used. Furthermore, the study location was one tertiary academic hospital. Further prospective validation across a variety of settings and multiple centers would significantly increase the credibility and generalizability of the results. The DL model's performance in a real-world screening scenario may differ from our controlled environment. Notwithstanding these limitations, the strengths of the study included a rigorous methodology and integration of the DL to the use of a commercially available POCUS system.

In conclusion, our study indicates that a DL-guided POCUS can serve as an effective tool for AAA screening, achieving diagnostic accuracy that is on par with experienced physicians. This innovative approach has the potential to democratize AAA screening, enhancing accessibility and cost-effectiveness. By harnessing the capabilities of AI, we can streamline the screening process, reduce the need for extensive sonographic training, and potentially improve patient outcomes through early detection.

Methods

Study design and ethical approval

This prospective evaluation study was conducted at the Kaohsiung Chang Gung Memorial Hospital in Taiwan between June and August 2023. Study approval was granted by the institutional review board at the hospital (IRB number: 202102311B0), and written consent was obtained from each participant.

Data collection and preprocessing

For developing the DL model, we collected ultrasound images from the ultrasound machines, include Sonosite Edge II and Hitachi Noblus, in the ED of Kaohsiung Chang Gung Hospital from January 2019 to December 2021. Ultrasound images focusing on the abdominal area were collected. Images that were not related to abdominal aorta examination were excluded, resulting in a dataset comprising 2101 labeled ultrasound images. This dataset was split into an 8:2 ratio for training and validation sets. We also collected 492 ultrasound images from a regional hospital for external validation.

As the original ultrasound images vary in size, we opted for the 'letterbox' preprocessing method to standardize them to the model's required dimensions. Letterboxing scales the image while maintaining the original aspect ratio; any remaining space after scaling is filled with the background, mimicking the effect of placing a picture into an envelope, hence the term 'letterbox'. Each ultrasound image was resized to 600 × 400 and get rid of the information that may reveal personal identification. Two medical experts on the point of care ultrasound then manually labeled the selected anatomical structures—the aorta, inferior vena cava (IVC), and spine (Supplemental Fig. 1) with a polygon mask. We adopted the commonly used labeling software, Labelme²⁵, for this study and saved the labeling file under COCO dataset format. This large, annotated dataset served as the foundation for training the AI models, enabling them to recognize and correctly identify these structures in ultrasound images.

Development of the DL model

We developed the DL model to offer real-time, continuous guidance during scanning to assist users in obtaining videos for AAA screening. The model, which emulates physician expertise, utilizes a You Only Look Once (YOLO) architecture, known for its real-time object detection capabilities. It is specifically tailored to analyze ultrasonographic images, focusing on identifying anatomical structures, including the abdominal aorta, spine, and inferior vena cava. The architecture employed in our study is YOLOv5 instance segmentation. The input is an ultrasound image, and the inference output includes bounding boxes and pixel area identification for each category (abdominal aorta, spine, and inferior vena cava). As the AI application scenario in this project is divided into two parts—real-time Aorta recognition for guidance and post-scan calculation of the maximum Aorta diameter—the former requires a faster model, while the latter requires a more accurate model. Therefore, among all YOLOv5 architectures, we trained YOLOv5s (225 layers, 7.4 million parameters, and approximately 26

GFLOPs of computation) for real-time guidance and YOLOv5m (367 layers, 21.2 million parameters, and approximately 73 GFLOPs of computation) for calculating the maximum Aorta diameter.

The training was conducted on an NVIDIA 3090 24GB GPU, utilizing the PyTorch framework. The training process we employed involved transfer learning from YOLOv5 pretrained on the COCO dataset, given that the ultrasound training data is relatively limited. We used the SGD optimizer with a batch size of 16 and a learning rate of 0.01 and trained the model for up to 100 epochs with early stopping if no improvement in validation performance was observed for 5 epochs. The result of validation is shown in Supplemental Table 1.

Integrated DL model with POCUS

The POCUS equipment used in this study is the ArtUs-EXT-1H from Teled, an FDA-certified platform for capturing raw ultrasound signals. It can be used in conjunction with a portable tablet computer. The tablet runs on a Windows 11 environment and uses an Intel CPU (detailed specifications are listed below). To accelerate the inference speed of the deployed model, the trained YOLOv5 model was converted from the PyTorch (.pt) format to the OpenVINO IR (FP16) format and utilized via OpenVINO Runtime. We used Python's built-in ctypes library to load the dll (dynamic link library) provided by Teled, allowing real-time ultrasound images to be captured within the Python program. The program also allows for model inference and uses OpenCV to visually represent the identified Aorta. After integration, the software monitors and processes the ultrasound display continuously through its application programming interface (Supplemental Fig. 2, Supplemental video 1). Additionally, to accommodate for the need to adjust parameters such as TGC (time gain compensation) during scanning, a control panel interface is displayed using Tkinter for the operator to adjust as necessary.

Prospective study design

Patients at least 65 years old visiting the outpatient clinic of the Cardiology department at the studied hospital were recruited between June and August 2023. Individuals were excluded if they were unable to lie flat or were unable or unwilling to provide informed consent.

Ten registered nurses without prior experience performing or interpreting ultrasonography were recruited for the trial from hospital personnel. Each nurse underwent a 15-min tutorial to familiarize him- or herself with the POCUS machine and DL guidance. Before undertaking the study, each nurse performed 1 practice scan on volunteer models to familiar with the software's user interface. They were instructed to acquire a 10-s standard abdominal aorta tracing video under DL guidance. For control, a duplicate scan was obtained by a physician using the same POCUS machine on the same day but without AI guidance. The physician also labeled the maximal width of the aorta for the control scan. The nurse scans were conducted independently, solely with DL guidance, and always preceded the control scans. Following each scan, the Teled POCUS machine stored two ultrasonography videos at 20 frames per second, which the DL system then processed to predict the maximal width of the abdominal aorta. Fig. 1 illustrates the study design.

Upon completion of all study and control scans, a panel of three expert physicians (Y.-C.Z., X.-H.L., and F.-J.C.) independently and blinded to whether a nurse or a physician performed the study, assessed whether each scan was of diagnostic quality, served as the primary endpoint. For cases with discrepancies, the majority rule was applied directly, where the judgment agreed upon by at least two experts was taken as the gold standard. All expert readers were certified board physicians in Cardiology or Emergency Medicine. The time to complete the study, defined as the interval from placing the probe on the patient's abdomen to completing the scan, was recorded. The maximal aortic width predicted by the DL model was compared with expert measurements for the secondary endpoint.

Statistical analysis

The study sought to evaluate the performance of nurses conducting AAA screening under DL guidance, with continuous variables reported as medians and interquartile ranges (IQR), and categorical variables as numbers and percentages. The proportion of qualified studies, as judged by the expert panel, was compared between DL guidance and physician scans for the primary endpoints using the non-inferiority test. This test provided a p -value to assess whether the DL-guided scans were not inferior to physician scans within a pre-specified margin. The maximal abdominal aortic width measurement and the time required to complete the study were evaluated as secondary endpoints and were compared using the Mann-Whitney U test. For both primary and secondary parameters, the proportion judged clinically evaluable is reported with 95% confidence intervals (CIs).

To measure the maximal width of the aorta from ultrasound video frames, we utilize the YOLOv5m architecture for bounding box prediction. Each video is 10 s in length with a frame rate of 20 frames per second (fps), resulting in a total of 200 frames per video. The model outputs bounding boxes, confidence scores, and class labels for detected objects in each frame. We adopted postprocessing steps, including non-maximum suppression²⁶, to filter out overlapping bounding boxes, ensuring that only the most confident predictions for the abdominal aorta are retained. We then extracted the four coordinates of the bounding box, which represent the top-left and bottom-right corners of the box enclosing the aorta. To calculate the width of the aorta, we averaged the differences between the coordinates on the vertical and horizontal axes. This process loops through each frame of the video, and the highest width, along with the corresponding image frame, is stored for human expert inspection.

Data availability

The raw data used in this study are not publicly available to preserve participant privacy.

Code availability

Code for data preprocessing, model training, and model inference is available. Refer to https://github.com/IMinChiu/AAA_guidance.

Received: 18 May 2024; Accepted: 23 September 2024;

Published online: 15 October 2024

References

- Roth, G. A. et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736–1788 (2018).
- Aggarwal, S., Qamar, A., Sharma, V. & Sharma, A. Abdominal aortic aneurysm: a comprehensive review. *Exp. Clin. Cardiol.* **16**, 11–15 (2011).
- Hofmann, B., Brandsaeter, I. Ø. & Kjelle, E. Variations in wait times for imaging services: a register-based study of self-reported wait times for specific examinations in Norway. *BMC Health Serv. Res.* **23**, 1287 (2023).
- Statistics., Statistics » *Diagnostics Waiting Times and Activity*, Retrieved (2024) from <https://www.england.nhs.uk/statistics/statistical-work-areas/diagnostics-waiting-times-and-activity/>.
- Benson, R. A., Meecham, L., Fisher, O. & Loftus, I. M. Ultrasound screening for abdominal aortic aneurysm: current practice, challenges and controversies. *Br. J. Radiol.* **91**, 20170306 (2018).
- Lee, L. & DeCara, J. M. Point-of-care ultrasound. *Curr. Cardiol. Rep.* **22**, 149 (2020).
- Taylor, R. A. & Moore, C. L. Point-of-care ultrasonography of the thoracic aorta. In *Emergency Point-of-Care Ultrasound* 32–38 (eds Connolly, J. A., Dean, A. J., Hoffmann, B. & Jarman, R. D.) (John Wiley & Sons, Ltd, Chichester, UK, 2017).
- Hashim, A. et al. The utility of point of care ultrasonography (POCUS). *Ann. Med. Surg. (Lond.)* **71**, 102982 (2021).

9. Dewar, Z. E. et al. A comparison of handheld ultrasound versus traditional ultrasound for acquisition of RUSH views in healthy volunteers. *J. Am. Coll. Emerg. Physicians Open* **1**, 1320–1325 (2020).
10. Britton, N. et al. Tele-ultrasound in resource-limited settings: a systematic review. *Front. Public Health* **7**, 244 (2019).
11. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **10**, 257–273 (2017).
12. Genovese, D. et al. Machine learning-based three-dimensional echocardiographic quantification of right ventricular size and function: validation against cardiac magnetic resonance. *J. Am. Soc. Echocardiogr.* **32**, 969–977 (2019).
13. Ouyang, D. et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* **580**, 252–256 (2020).
14. Narang, A. et al. Utility of a deep-learning algorithm to guide novices to acquire echocardiograms for limited diagnostic use. *JAMA Cardiol.* **6**, 624–632 (2021).
15. Cheng, C.-Y. et al. Deep learning assisted detection of abdominal free fluid in Morison's pouch during focused assessment with sonography in trauma. *Front. Med. (Lausanne)* **8**, 707437 (2021).
16. Chiu, I.-M. et al. Use of a deep-learning algorithm to guide novices in performing focused assessment with sonography in trauma. *JAMA Netw. Open* **6**, e235102 (2023).
17. Hartshorne, T. C., McCollum, C. N., Earnshaw, J. J., Morris, J. & Nasim, A. Ultrasound measurement of aortic diameter in a national screening programme. *Eur. J. Vasc. Endovasc. Surg.* **42**, 195–199 (2011).
18. Becker, D. M. et al. The use of portable ultrasound devices in low- and middle-income countries: a systematic review of the literature. *Trop. Med. Int. Health* **21**, 294–311 (2016).
19. Shaddock, L. & Smith, T. Potential for use of portable ultrasound devices in rural and remote settings in Australia and other developed countries: a systematic review. *J. Multidiscip. Healthc.* **15**, 605–625 (2022).
20. Thompson, S. G., Ashton, H. A., Gao, L., Buxton, M. J., Scott, R. A. & Multicentre Aneurysm Screening Study (MASS) Group. Final follow-up of the multicentre aneurysm screening study (MASS) randomized trial of abdominal aortic aneurysm screening. *Br. J. Surg.* **99**, 1649–1656 (2012).
21. McCaul, K. A., Lawrence-Brown, M., Dickinson, J. A. & Norman, P. E. Long-term outcomes of the western Australian trial of screening for abdominal aortic aneurysms: secondary analysis of a randomized clinical trial. *JAMA Intern. Med.* **176**, 1761–1767 (2016).
22. Madani, A., Ong, J. R., Tibrewal, A. & Mofrad, M. R. K. Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. *NPJ Digit. Med.* **1**, 59 (2018).
23. Duffy, G. et al. High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *JAMA Cardiol.* **7**, 386–395 (2022).
24. Cheng, C.-Y. et al. Development and validation of a deep learning pipeline to measure pericardial effusion in echocardiography. *Front. Cardiovasc. Med.* **10**, 1195235 (2023).
25. LabelMe. *The Open Annotation Tool*. Retrieved (2022). <http://labelme.csail.mit.edu>.
26. Hosang, J., Benenson, R. & Schiele, B. Learning non-maximum suppression. In proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017).

Acknowledgements

This research was funded by the National Science and Technology Council (MOST 111-2622-E-182A 001-). The funders played no role in the study design, data collection, data analysis, data interpretation, or writing of this manuscript.

Author contributions

I.-M.C. and C.-Y.C. conceptualized the project, managed and coordinated the project, assisted with the design of the methodology, analyzed the data, and prepared the initial and final drafts of the manuscript. I.-M.C., D.O., and C.-Y.C. analyzed the data and prepared the initial and final drafts of the manuscript. Y.-C.Z., X.-H.L., and F.-J.C. take responsibility for the data and their analysis. T.-Y.C. contributed to the management and coordination of the project, assisted with the design of the methodology, and helped review the manuscript. All the authors performed a critical review and approved the final manuscript for interpretation of the data and important intellectual input.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-024-01269-4>.

Correspondence and requests for materials should be addressed to David Ouyang or Chi-Yung Cheng.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024