



# Evaluating the benefits of machine learning for diagnosing deep vein thrombosis compared with gold standard ultrasound: a feasibility study

Kerstin Nothnagel<sup>1\*</sup>, Mohammed Farid Aslam<sup>2</sup>

<sup>1</sup>Population Health Sciences, Canynge Hall, Bristol Medical School, University of Bristol, Bristol, UK; <sup>2</sup>Imperial College London, London, England, UK

## Abstract

**Background:** This study evaluates the feasibility of remote deep venous thrombosis (DVT) diagnosis via ultrasound sequences facilitated by ThinkSono Guidance, an artificial intelligence (AI) app for point-of-care ultrasound (POCUS).

**Aim:** To assess the effectiveness of AI-guided POCUS conducted by non-specialists in capturing valid ultrasound images for remote diagnosis of DVT.

**Design & setting:** Over a 3.5-month period, patients with suspected DVT underwent AI-guided POCUS conducted by non-specialists using a handheld ultrasound probe connected to the app. These ultrasound sequences were uploaded to a cloud dashboard for remote specialist review. Additionally, participants received formal DVT scans.

**Method:** Patients underwent AI-guided POCUS using handheld probes connected to the AI app, followed by formal DVT scans. Ultrasound sequences acquired during the AI-guided scan were uploaded to a cloud dashboard for remote specialist review, where image quality was assessed, and diagnoses were provided.

**Results:** Among 91 predominantly older female participants, 18% of scans were incomplete. Of the rest, 91% had sufficient quality, with 64% categorised by remote clinicians as 'compressible' or 'incompressible'. Sensitivity and specificity for adequately imaged scans were 100% and 91%, respectively. Notably, 53% were low risk, potentially obviating formal scans.

**Conclusion:** ThinkSono Guidance effectively directed non-specialists, streamlining DVT diagnosis and treatment. It may reduce the need for formal scans, particularly with negative findings, and extend diagnostic capabilities to primary care. The study highlights AI-assisted POCUS potential in improving DVT assessment.

\*For correspondence: Kerstin. Nothnagel@bristol.ac.uk

**Competing interest:** The authors declare that no competing interests exist.

**Received:** 02 March 2024

**Accepted:** 15 May 2024

**Published:** 18 December 2024

©This article is Open Access: CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

**Author Keywords:** clinical (general), screening, diagnosis, venous thrombosis, artificial intelligence

Copyright © 2024, The Authors; DOI:10.3399/BJGPO.2024.0057

## How this fits in

As the global population continues to age, the prevalence of deep venous thrombosis (DVT) is on the rise. Presently, DVT diagnostic protocols necessitate specialised equipment operated by trained personnel. The advent of artificial intelligence (AI)-guided point-of-care ultrasound (POCUS) offers the potential for non-specialists to conduct DVT scans. If proven accurate and feasible, the implementation of AI-guided POCUS could alleviate pressure on secondary care services, enhance DVT diagnostic availability in primary care, and contribute to greater efficiency within the NHS, ultimately reducing diagnostic expenses.

## Introduction

The National Institute for Health and Care Excellence (NICE) reports that deep vein thrombosis (DVT) has an annual incidence of 1–2 per 1000 people.<sup>1</sup> Ensuring an accurate and prompt diagnosis is vital to forestall complications such as pulmonary embolism (PE) or post-thrombotic syndrome (PTS).<sup>2</sup> Although progress has been made, DVT diagnosis is embedded in a complex clinical pathway incorporating risk assessment, laboratory analyses, and ultrasound, the latter typically performed in secondary care. The Wells score categorises patients into low, intermediate, and high risk, with a subsequent D-dimer blood test often used to rule out DVT in patients who are intermediate risk, despite its low specificity.<sup>3</sup> Nevertheless, the traditional approach involves utilising compression ultrasound in conjunction with Doppler ultrasound (duplex). This method involves observing intraluminal contents and increases accuracy by revealing potential lack of blood flow where thrombosis is present but not visible.<sup>4</sup>

Previous studies have demonstrated that POCUS for DVT diagnosis, even without AI support, exhibits high sensitivity and specificity.<sup>5</sup> For example, this has been shown in the meta-analysis conducted by Lee *et al*,<sup>5</sup> in the randomised controlled trial by Bernardi *et al*,<sup>6</sup> and the systematic review and meta-analysis by Kraaijpoel *et al*,<sup>7</sup> which concluded that POCUS of the proximal leg yields results comparable with standard DVT diagnostic ultrasound.

The need for more accessible diagnostic methods is evident, and the integration of AI-guided POCUS in primary care holds promise to address these challenges. Ultrasound effectiveness remains contingent on user proficiency. Although studies suggest that less trained healthcare professionals (HCPs) can reliably perform POCUS examinations,<sup>4</sup> specific training requirements remain unclear.<sup>8</sup> Consequently, highly trained specialists, including radiologists, and ultrasound technicians, are often required for accurate documentation,<sup>9</sup> and liability mitigation, incurring significant costs, and longer waiting times.<sup>10</sup> AI-guided POCUS has the potential to reduce the need for secondary care admissions and enhance access to DVT diagnostics in primary care settings. Particularly, it would benefit the most vulnerable patient groups such as the housebound, frail, and multimorbid. The implementation of this technology aligns with the goals outlined in the NHS England's long-term plan, which envisions to improve access to diagnostics in primary care through use of innovative technologies.<sup>11</sup> The aim of this feasibility study was to:

1. Evaluate AI-guided POCUS effectiveness in guiding HCPs without formal ultrasound training to obtain ultrasound sequences of sufficient quality for remote interpretation.
2. Investigate if remote specialists can accurately provide a diagnosis based on AI-guided POCUS acquired sequences, comparing their assessments with those from the reference scan.
3. Assess the efficiency as a triage tool, stratifying scans into high and low risk and potentially eliminating the need for additional scans by a specialist, thereby streamlining the diagnostic process.

## Method

### Patient recruitment

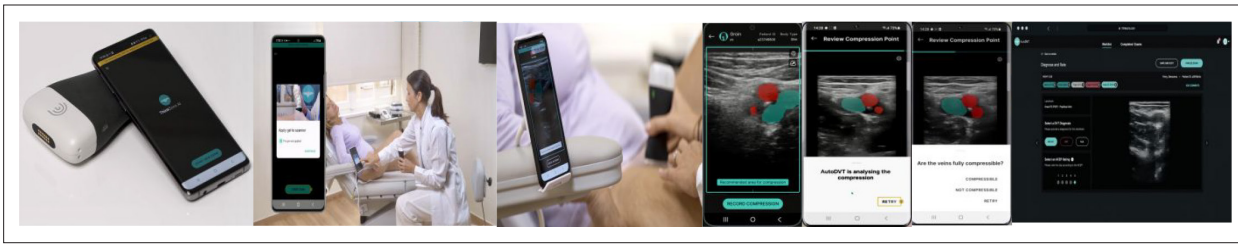
Patients with suspected DVT requiring an ultrasound scan (USS) were recruited consecutively at a hospital in Berlin, Germany, over a 3.5-month period in 2022. Forty-six per cent of the participants were inpatients, 2% had directly scheduled scan appointments, and the majority (52%) were referred from the primary care sector.

A statistical dichotomous endpoint power calculation was conducted by an Imperial College London statistician. The calculation was based on the understanding that approximately 15% of patients with symptoms indicative of DVT would test positive. To achieve a statistical power of 80%, a minimum of seven patients with positive DVT cases were needed to be included in the study.

Patients were eligible for inclusion if they were aged  $\geq 18$  years, demonstrated the capacity to consent, exhibited symptoms suggestive of DVT, and the DVT diagnostic algorithm indicated the necessity of a diagnostic USS.<sup>12</sup>

Each consenting patient underwent two distinct USS:

- index: AI-guided POCUS conducted by HCPs without formal DVT diagnostic ultrasound training; and
- reference: the second scan, a formal duplex scan, which was conducted by a specialist blinded to the index scan results.



**Figure 1** A–G Workflow of ThinkSono Guidance (it guides the operator through anatomical landmarks and indicates when to compress. ThinkSono Guidance analyses the images, and provides a remote diagnosis)

## Index test

The AI-guided scans were conducted using ThinkSono Guidance, an app developed by the company ThinkSono GmbH. This app, installed on a smartphone, was connected to a Clarius L7 HD3 linear handheld ultrasound probe. The app had, at the time of the study, a Class 1 CE certificate (Supplementary Information S1) and was used under its intended purpose (Supplementary Information S2).

Prior validation of the ThinkSono Guidance has been documented in previous studies.<sup>6,7</sup> The app guides users through a two-region POCUS, directing compressions in the groin areas and in the knee pit. If the procedure identifies insufficient compression, improper vessel alignment in the image, or probe displacement during compression, the application suggests redoing the compression. Users retain the choice to override these suggestions. After successfully gathering all compression points, sequences of compression (lasting 10 seconds each) are uploaded to a cloud-based dashboard for remote assessment.

Four sequences are recorded per scan (**Figure 1a–g**). USS were conducted by a non-specialist without formal experience in performing DVT diagnostic scans, following a 1-hour training session with ThinkSono Guidance.

## AI algorithm

The ThinkSono Guidance app utilises a convolutional neural network for vessel analysis, supplemented by auxiliary branches for anatomical location prediction. Trained on prospectively collected data, the model employs manual vein and artery delineations along with image-level labels for anatomical locations. It operates on B-mode ultrasound images resampled to  $128 \times 128$  pixels, producing segmentation masks and categorical location labels within 25 milliseconds.

The U-Net architecture serves as the model backbone, augmented by a landmark prediction branch. Three models, each tailored to specific anatomical regions, are trained with identical architecture. Further details and illustrations are provided in the figures and tables within the article.

The U-Net architecture serves as the model backbone, augmented by a landmark prediction branch. Three models, each tailored to specific anatomical regions, are trained with identical architecture. Further details and illustrations are provided in the attached figures.

## Reference standard

The DVT diagnostic scan, recommended by NICE guidelines, includes compressions at multiple points on the proximal leg, coupled with duplex.<sup>12</sup> Compressions are typically applied every few centimetres from the groin to the knee pit along the large veins. Duplex is employed on the large veins to demonstrate normal blood movement. A qualified physician, unaware of the index scan outcomes, provided a report after completion of the reference scans, indicating the presence or absence of DVT.

## Image review

The ultrasound sequences from the index scan were assessed by five qualified remote experts using the secure cloud dashboard. These experts were kept unaware of all patient attributes and indications, only being informed if the examination was conducted on the left or right extremity.

Remote specialists assessed image quality using the American College of Emergency Physicians (ACEP) score, which is rated on a scale from 1–5. An overall examination score of  $\geq 3$  was deemed to indicate sufficient image quality (**Table 1**). Furthermore, each image scoring  $\geq 3$  was categorised as having 'compressible' veins, 'incompressible' veins, or requiring another USS ('indeterminate').

**Table 1** American College of Emergency Physicians (ACEP) score

ACEP score	
1	No recognisable structures, no objective data can be gathered
2	Minimally recognisable structures, but insufficient for diagnosis
3	Minimal criteria met for diagnosis, recognisable structures but with some technical or other flaws
4	Minimal criteria met for diagnosis, all structures images well and diagnosis supported easily
5	Minimal criteria met for diagnosis, all structures images with excellent image quality and diagnosis completely supported

## Patient triaging

To mimic clinical triage methods, a retrospective analysis combined the classifications into high-risk and low-risk groups. This process aimed to reflect the proportion of scans that could potentially be managed in primary care without the need for referral to a formal USS. The scans were evaluated for image quality and compressibility assessment, as summarised in **Table 2**.

## Statistical analysis

This study sought to assess the feasibility of utilising AI-guided POCUS to guide healthcare professionals (HCPs) without formal ultrasound training in acquiring ultrasound images suitable for remote interpretation. The evaluation included an analysis of image quality, the accuracy of remote diagnoses, and the efficiency of the AI-guided POCUS application as a primary care triage tool by categorising scans into high- and low-risk groups, potentially reducing the necessity for scans in secondary care.

For the reported confidence intervals (95% CI), statistical methods recommended for diagnostic test accuracy analysis were applied. This approach follows standard practices used in previous studies, including the development phase of ThinkSono Guidance (formerly AutoDVT), as outlined in the study by Prof. Kainz<sup>13</sup>. These methods were used to calculate the sensitivity and specificity of the AI-guided POCUS, contributing to the overall assessment of its diagnostic performance.

The analysis was conducted using Google Sheets, and the outcomes were displayed with variability divergence. The image quality assessment employed a majority voting process with five experts assigning an ACEP score of  $\geq 3$  for the uploaded scans. Remote diagnoses relied on the unanimous agreement of at least three experts regarding vein compressibility, and the results were compared with formal USS, considering sensitivity and specificity. The triaging into low- and high-risk categories was performed through majority voting by the remote specialists.

For the reported confidence intervals (95% CI), statistical methods recommended for diagnostic test accuracy analysis were applied. This approach follows standard practices used in previous studies, including the development phase of ThinkSono Guidance (formerly AutoDVT)<sup>13</sup>. These methods were used to calculate the sensitivity and specificity of the AI-guided POCUS, contributing to the overall assessment of its diagnostic performance.

## Results

### Participants

The 91 participants included 59% females, with an average age of 69.7 years. The mean body mass index (BMI) was calculated at 26, with the highest BMI recorded at 47.3 and the lowest at 15.8. It is noteworthy that only 59% of the 91 participants were referred for a scan based on a genuine suspicion of

**Table 2** Remote specialist reported diagnosis, index scan and sonographer diagnosis, and reference scan

	Reference scan positive	Reference scan negative
Index scan positive	9% TP	9% FP
Index scan negative	0% FN	83% TN

FN = false negative. FP = false positive. TN = true negative. TP = true positive.

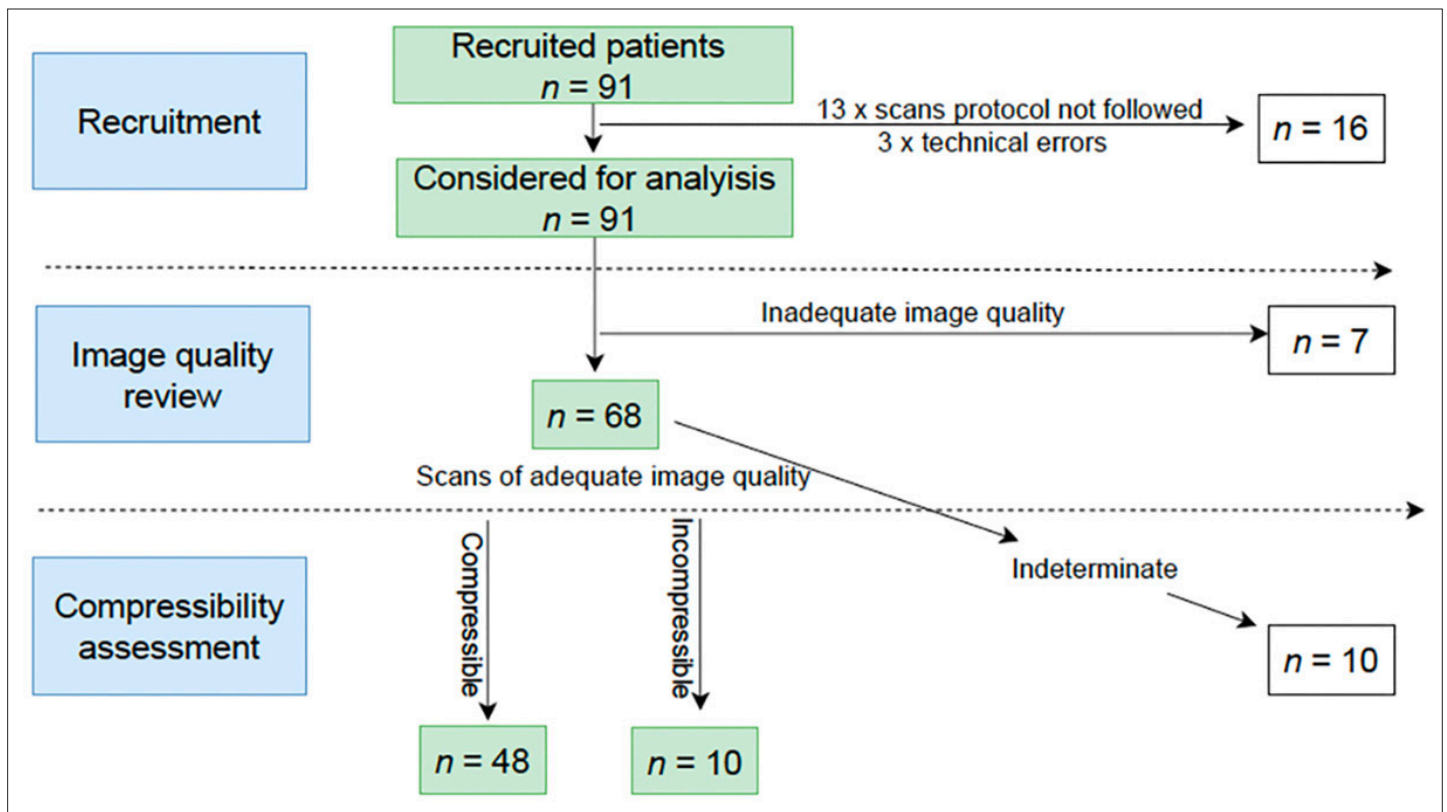


Figure 2 Flowchart

DVT. The remaining 41% of scans were conducted as exclusion diagnostics before initiating treatment for conditions such as cellulitis with antibiotics or for musculoskeletal issues requiring physiotherapy. This highlights that a majority of diagnostics, especially in secondary care, are performed as exclusion diagnostics.

Thirteen of the 91 scans deviated from the study protocol, and an additional three encountered technical errors. These USS (n = 16, 18%) were either fragmented or not correctly transferred to the cloud dashboard, preventing an overall diagnosis (Figure 2). Consequently, these scans were categorised as high-risk group. Out of the 75 uploaded scans, 91% achieved an ACEP score of  $\geq 3$ , indicating sufficient image quality for remote diagnosis as highlighted in Table 2.

However, in 15% of these scans, despite adequate image quality, a remote diagnosis could not be provided owing to issues such as other pathological findings, for example, a Baker's cyst or disagreement between the reviewers.

Table 3 Summary of key information

All scans	Result
Image quality <sup>a</sup>	75% (68/91) scans with ACEP $\geq 3$
Remote diagnosis	64% (58/91) scans included for diagnosis
Risk triaging <sup>b</sup>	53% (48/91) of participants at low risk
<b>Scans considered for remote analysis</b>	
Image quality <sup>a</sup>	91% (68/75) scans with ACEP $\geq 3$
Remote diagnosis	77% (58/75) scans included for diagnosis
Risk triaging <sup>b</sup>	64% (48/75) of participants at low risk

<sup>a</sup>Demonstrating adequate quality for remote interpretation. <sup>b</sup>Suggesting potential avoidance of a formal scan in secondary care

## Image quality

In the process of collective decision making, the five remote specialists allocated an ACEP score of  $\geq 3$  for 68 out of the 75 examinations, while seven examinations received an ACEP score  $< 3$ . The average ACEP score for these exams was 3.35. Consequently, 75% (68/91) of the scans demonstrated adequate quality for diagnosis by remote clinicians (Table 4).

## Remote diagnosis

**Table 2** Of the ultrasound scans reviewed by remote clinicians, 64% were categorised as either 'compressible' or 'incompressible,' aligning with the primary study objective of evaluating diagnostic feasibility (Table 4). However, 15% of the 68 scans encountered challenges that impeded a thorough assessment, including findings of other pathologies or disagreements among reviewers. Among the remaining 58 scans, the outcomes were as follows: 9% true positive (TP), 9% false positive (FP), 83% true negative (TN), and 0% false negative (FN), as illustrated in Table 3. This resulted in a sensitivity of 100% (95% confidence interval [CI] = 99.12% to 100%), indicating its ability to correctly identify all true positive cases, and a specificity of 90.57% (95% CI = 90.48% to 91.66%), highlighting its capability to accurately rule out the presence of DVT in the majority of cases. These findings suggest a high level of reliability in both sensitivity and specificity for AI-guided POCUS in diagnosing DVT.

## Risk triaging

If AI-guided USS with remote diagnosis would be employed as triage, 48% of 91 USS included in the study would have been categorised as high risk: 18% of 91 scans deviated from the study protocol, 8% had an ACEP score  $< 3$ . Additionally, 11% were classified as needing to be repeated ('indeterminate') despite sufficient image quality and 11% of scans were categorised as 'incompressible'. In contrast, the remaining 53% of scans were marked as 'low risk', as each sequence displayed compressible veins with adequate image quality. In clinical settings, patients at low risk would potentially not require a formal USS (Table 3).

## Discussion

### Summary

This study suggests that AI-assisted DVT diagnosis conducted by HCPs without formal ultrasound training can effectively capture quality images of proximal leg veins using AI-guided POCUS. This method shows promise for remote diagnosis and integration into primary care, potentially reducing the demand for formal ultrasounds, thereby enhancing diagnostic efficiency and resulting in cost savings.

### Strengths and limitations

While our study demonstrates promising outcomes, it has limitations. The statistical analysis was retrospective, and further prospective research is needed to validate the method's safety and effectiveness. Additionally, the study had a small number of positive DVT cases, indicating the

**Table 4** Risk triaging

Risk category	Criteria for designation
High risk	Incomplete examination
	At least 1 sequence with ACEP $< 3$
	Incompressible veins
	Other reasons deemed by reviewer, for example, considerations for differential diagnosis such as a Baker's Cyst
Low risk	Complete examination
	All cine-loop sequences with ACEP $\geq 3$
	All veins are identified as compressible

ACEP = American College of Emergency Physicians

necessity for validation with larger cohorts. Moreover, logistical constraints prevented real-time feedback, potentially affecting the ratio of avoided formal ultrasounds.

## Comparison with existing literature

Our findings align with previous research on the reliability of POCUS for DVT diagnosis.<sup>6,7</sup> Integrating AI technology has the potential to enhance POCUS accuracy further. However, there are concerns regarding documentation and liability when scans are conducted by non-experts. Our method addresses some of these concerns by integrating remote expert review and recommending a formal scan if suspicion for DVT remains high..

## Implications for research and practice

Integrating AI-guided DVT diagnosis into primary care could revolutionise diagnostic processes and extend diagnostic capabilities to underserved populations.<sup>6</sup> This technology could streamline diagnostic pathways, particularly benefiting housebound or nursing home residents. Further research is needed to validate our findings in larger, multi-centre studies and to assess the generalisability of our results to broader healthcare contexts.

In conclusion, our findings support the feasibility of AI-guided POCUS with remote expert diagnosis. It could potentially offer a streamlined and cost-effective approach with the potential for improved patient outcomes.

### Funding

This study received no external funding. The sponsor, Imperial College London, covered costs related to the study, including the engagement of a statistician and the processing of ethical applications.

### Ethical approval

Ethical approval for this study was obtained from the German ethics committee 'Ärzttekammer Berlin' with the approval number ETH-07/21. Furthermore, the study had to be thoroughly inspected by the Data Protection Impact Assessment (DPIA).

### Trial registration number

ISRCTN24293748

### Provenance

Freely submitted; externally peer reviewed.

### Data

The dataset relied on in this article is available from the corresponding author on reasonable request.

### Acknowledgements

I would like to express my gratitude to my PhD supervisors, Dr Watson, and Prof Hay, for providing valuable insights and assistance in crafting this publication. Their input has been invaluable, contributing to the quality and coherence of the manuscript.

## References

1. National Institute for Health and Care Excellence (NICE). *Deep vein thrombosis: what are the risk factors?* 2023. <https://cks.nice.org.uk/topics/deep-vein-thrombosis/background-information/risk-factors/> (accessed 28 Oct 2024).
2. Lewiss RE, Kaban NL, Saul T. Point-of-care ultrasound for a deep venous thrombosis. *Glob Heart* 2013; **8**(4): S2211-8160(13)00164-6): 329–333: . DOI: <https://doi.org/10.1016/j.gheart.2013.11.002>
3. Bhatt M, Braun C, Patel P, et al. Diagnosis of deep vein thrombosis of the lower extremity: a systematic review and meta-analysis of test accuracy. *Blood Adv* 2020; **4**(7): 1250–1264. DOI: <https://doi.org/10.1182/bloodadvances.2019000960>
4. Zitek JA, Baydoun J, Baird J. Tools for the clinician: the essentials of bedside (ED or ICU) ultrasound for deep vein thrombosis. *Curr Emerg Hosp Med Rep* 2013; **1**(2): 65–70. DOI: <https://doi.org/10.1007/s40138-013-0016-4>

5. Lee JH, Lee SH, Yun SJ. Comparison of 2-point and 3-point point-of-care ultrasound techniques for deep vein thrombosis at the emergency department: a meta-analysis. *Medicine (Balt)* 2019; **98**(22): e15791. DOI: <https://doi.org/10.1097/MD.00000000000015791>
6. Bernardi E, Camporese G, Büller HR, et al. Serial 2-point ultrasonography plus D-dimer vs whole-leg color-coded Doppler ultrasonography for diagnosing suspected symptomatic deep vein thrombosis: a randomized controlled trial. *JAMA* 2008; **300**(14): 1653–1659. DOI: <https://doi.org/10.1001/jama.300.14.1653>
7. Kraaijpoel N, Carrier M, Le Gal G, et al. Diagnostic accuracy of three ultrasonography strategies for deep vein thrombosis of the lower extremity: a systematic review and meta-analysis. *PLoS One* 2020; **15**(2): e0228788. DOI: <https://doi.org/10.1371/journal.pone.0228788>
8. Akanuwe JNA, Siriwardena AN, Bidaut L, et al. Practitioners' views on community implementation of point-of-care ultrasound (POCUS) in the UK: A qualitative interview study. *BMC Health Serv Res* 2023; **23**(1): 84. DOI: <https://doi.org/10.1186/s12913-023-09069-4>
9. Fischer EA, Kinnear B, Sall D, et al. Hospitalist-operated compression ultrasonography: a point-of-care ultrasound study (HOCUS-POCUS). *J Gen Intern Med* 2019; **34**(10): 2062–2067. DOI: <https://doi.org/10.1007/s11606-019-05120-5>
10. Heerink JS, Nies J, Koffijberg H, et al. Two point-of-care test-based approaches for the exclusion of deep vein thrombosis in general practice: a cost-effectiveness analysis. *BMC Prim Care* 2023; **24**(1): 42. DOI: <https://doi.org/10.1186/s12875-023-01992-z>
11. NHS England. *NHS Long Term Plan* 2019. <https://www.england.nhs.uk/publication/the-nhs-long-term-plan/> (accessed 28 Oct 2024).
12. NICE. *Quality statement 3: Proximal leg vein ultrasound scan for a "likely" deep vein thrombosis Wells score* 2021. <https://www.nice.org.uk/guidance/qs201/chapter/Quality-statement-3-Proximal-leg-vein-ultrasound-scan-for-a-likely-deep-vein-thrombosis-Wells-score#:~:text=Rationale> (accessed 28 Oct 2024).
13. Kainz B, Heinrich MP, Makropoulos A, Oppenheimer J, et al. Non-invasive diagnosis of deep vein thrombosis from ultrasound imaging with machine learning. *NPJ Dig Med* 2021; **4**(1): 137: . DOI: <https://doi.org/10.1038/s41746-021-00503-7>