

Artificial Intelligence as a tool for diagnostic support in chronic venous disease

M Pizzamiglio^{1,2}, DE Abad Gomez¹, P Casoni, E Nanni², F Álvarez Marcos³, F Gallardo Pedrajas¹

¹Department of Vascular Surgery, Hospital Quironsalud Marbella, Av. Severo Ochoa, 22, 29603 Marbella, Málaga, Spain. Email: mpizzamiglio.3@gmail.com. Phone: 0034686596560.

²Ippocrate Vein Clinic, Parma, Italy.

³Department of Vascular Surgery, Asturias University Central Hospital (HUCA), Oviedo, Spain.

submitted: May 5, 2025, accepted: Jun 19, 2025, Epub Ahead of Print: Aug 31, 2025, published: Aug 31, 2025
Conflict of interest: None

DOI: [10.24019/jtavr.234](https://doi.org/10.24019/jtavr.234) - Corresponding author: Dr. Matteo Pizzamiglio, mpizzamiglio.3@gmail.com

© 2025 Fondazione Vasculab impresa sociale ETS. All rights reserved.

Keywords Artificial Intelligence, Chronic Venous Disease, Clinical Decision Support, Diagnosis, Computer-Assisted, Natural language Processing.

Abstract This retrospective study aimed to assess the diagnostic accuracy of an AI-powered clinical decision support platform in diagnosing chronic venous disease (CVD), using specialist diagnosis as the reference. A total of 130 patients who presented to a vascular surgery consultation in January 2024 with both venous and non-venous pathologies were included. Clinical data, including patient histories, physical examination findings, and complementary test results, were manually entered into the AI platform. The AI-generated diagnoses were then compared to those made by vascular specialists using a standardized classification system. The level of agreement between the AI and the specialists was evaluated using Fleiss' kappa coefficient. The AI platform demonstrated a diagnostic accuracy of 97% for CVD, with a very high concordance (kappa = 0.925, 95% CI 0.81-1.03, $p < 0.001$). While the platform demonstrated excellent diagnostic accuracy for CVD, the limited representation of other pathologies in the sample limited the generalizability of the findings. The results suggest that while the AI platform shows promise as a tool to support clinical decision-making in CVD, further research is necessary to determine its effectiveness in broader diagnostic contexts.

Accurate and detailed data input remains essential to maximize the platform's diagnostic capabilities.

Introduction

Artificial Intelligence (AI) is a field of computer science that aims to mimic human thought processes, learning capacity, and knowledge storage¹. Healthcare, as in many other sectors, is living a rapid and transformative integration with artificial intelligence. It has been applied already to many specialties, such as Radiology², Cardiology³, Anesthesiology⁴ and even Angiology and Vascular Surgery⁵ among others. As we live this inevitable integration with AI, it's key for healthcare professionals to understand how it works and how it can be used to offer a more safe, efficient and cost-effective attention⁶.

Artificial intelligence embarks many areas, which include Machine Learning (ML), natural language processing (NLP) and Computer Vision. ML is defined as the ability to learn and make predictions based on pattern recognition. It has the ability to recognize tendencies, patterns and associations with large database. Deep Learning is a type ML that uses artificial neural networks that basically imitate human brain. It can learn from large databases, improve with every interaction and

complete complex tasks without depending on human intervention. Computer Vision is the ability to process images and videos. NLP is the capacity to process human language in text form and is useful to analyze big database such as clinical histories or radiological reports. Large Language Models are an advanced subgenre of NLP created to understand human text and able to work with it in a large scale, being able to generate a coherent text in an appropriate context from integrated prompts⁵.

Chronic venous disease is a highly prevalent condition⁸. It's diagnosis based on clinical history, physical examination and Doppler ultrasound is usually

straightforward. However, patients may present with overlapping conditions or complex symptoms that challenge the clinician's ability to establish a definitive diagnosis and corresponding treatment plan.

The objective of this study is to analyze the diagnostic precision of an AI-powered Clinical decision support platform (Glass Health AI; <https://glass.health>)⁷ for the diagnosis of chronic venous disease.

All patients included in this study provided informed consent for the use of their anonymized clinical data and images for research and publication purposes.

Material and Methods

During the month of January of 2024, patients who presented to consultation of vascular surgery for venous and non-venous pathologies were retrospectively and consecutively included. Only first consultation visits were included. Follow-up visits were excluded. Patient history, physical examination and results of doppler imaging were collected. These data were included in the artificial platform. For every patient, the AI platform gave a differential diagnosis. The disease ranked first was taken as the main diagnosis. The diagnostic precision given by the platform was evaluated comparing it to the diagnosis given by the specialist.

Clinical histories, physical examination findings and results from Vascular Doppler Ultrasound were collected and then entered into the platform individually for each patient. Chronic venous disease (CVD) was defined as "(any) morphological and functional abnormalities of the venous system of long duration manifest either by symptoms and/or signs indicating the need for investigation and/or care"⁹. Its diagnosis was established based on the CEAP classification (a standardized system that categorizes chronic venous disorders according to Clinical signs, Etiology, Anatomy, and Pathophysiology), taking into account clinical signs, patient-reported symptoms, and Doppler ultrasound findings.

Based on the differential diagnoses generated by the AI, each diagnostic option was mapped to a standardized classification system comprising the following categories:

1. Chronic Venous Disease, 2. Pelvic Venous Disease, 3. Post-thrombotic syndrome, 4. Peripheral neuropathy, 5. Peripheral arteriopathy, 6. Pseudoaneurysm, 7. Thromboangiitis obliterans, 8. Complication after hip surgery, 9. Chronic mesenteric ischemia, 10. Critical limb ischemia, 11. Superficial venous thrombosis, 12. Cellulitis, 13. Raynaud's phenomenon, 14. Endometriosis.

This protocol ensured that all input data adhered to a consistent structure, reducing variability and improving reliability in the analysis.

The variables "Human Diagnosis" and "AI Diagnosis" were recoded based on this common scale to ensure uniformity in the comparison. Fleiss' Kappa coefficient was calculated to measure the degree of agreement between the specialist and the artificial intelligence platform for the diagnosis, considering the global concordance and for each category. This coefficient was chosen for its capacity to evaluate the consistency between multiple observers in categoric scales, while accounting for agreement that could occur by chance.

Chronic venous disease (CVD), pelvic congestion syndrome (PCS), and post-thrombotic syndrome (PTS) are distinct clinical manifestations within the spectrum of venous pathology. Although they share overlapping pathophysiological mechanisms, each condition presents with unique clinical features, diagnostic criteria, and therapeutic approaches. For the purposes of this study, these entities were classified and analyzed separately in order to better define the diagnostic performance of the AI platform for each condition.

During this study, patient data was managed rigorously to ensure privacy and confidentiality. It was collected retrospectively and anonymized before analysis to protect the patient privacy. Additionally, only the data necessary to meet the study objectives were included (clinical history, physical examination and vascular Doppler Ultrasound). Since the data analysis was performed retrospectively and solely for research purposes, informed consent was not required.

Approval for this specific study was waived by local Institutional Review Board, according to the National Policy in the matter of the Privacy Act on retrospective analysis of anonymized data.

Results

A total of 320 patients were visited during the month of January of 2024. After excluding the follow ups, 130 patients were included for the study. 89 were women and 41 men. Mean age was 54.85 +/- 16.05

years. 18% of patients presented with high arterial blood pressure, 11% had Diabetes, 15% hyperlipidemia. Six patients had cardiologic history (coronary bypass and mitral insufficiency). Three patients had respiratory history (asthma and bronchiectasis). One patient had chronic renal insufficiency, 10% had thyroid affection (thyroid cancer treated surgically, Hashimoto thyroiditis and hypothyroidism). Three patients had gynecologic history (one with uterus cancer treated by total hysterectomy and two had polycystic ovarian). Four patients had lumbar discopathy. Two patients had history of venous thromboembolism (femoropopliteal DVT and PTE postpartum).

The diagnostic accuracy of the AI platform compared to the diagnosis made by the specialists for chronic venous disease was 97%, with a Kappa coefficient of 0.925 (IC 95% 0.81-1.03, $p<0.001$). This indicates a very high grade of concordance. The analysis for individual categories also showed a very high performance for the diagnosis of CVD, with a Kappa coefficient of 0.985 (IC 95% 0,76-1.11, $p<0.001$). The performance of diagnostic accuracy was poor for the other diagnostic categories, such as post-thrombotic syndrome and peripheral neuropathy. This isn't surprising because the sample and the goal of the study weren't done for the study of these diseases.

Cases	Age	Sex	Specialist diagnosis	AI diagnosis	Observations
1	35	F	Chronic venous insufficiency	Pelvic Congestion Syndrome	Relevant obstetric history
2	66	M	Chronic venous insufficiency	Peripheral Neuropathy	Nonspecific symptoms, insufficient clinical data
3	34	F	Chronic venous insufficiency	Post-thrombotic Syndrome	History of pulmonary thromboembolism

Table 1. Cases of diagnostic discordance between human and AI diagnosis

Three cases showed diagnostic discordance between the platform and the vascular specialist (Table 1). The first was a female patient who presented to consultation with symptoms of heaviness of the lower limbs that begun after her second pregnancy. In the physical exploration small reticular veins were objectivated. The doppler echography didn't reveal any deep neither superficial reflux (the reverse flow of blood in veins detected via Doppler ultrasound) in standing position. They main diagnosis suggested by the AI was pelvic venous disease. The second patient was a 66 year old male who presented with paresthesia in both legs. The doppler echography showed insufficiency of the great saphenous vein of the left leg, being sufficient in the thigh. The rest of the saphenous axis were permeable and sufficient with distal compression and Valsalva maneuver. The main diagnosis of the AI was peripheral neuropathy. The third case was the one of a woman of 34 years old with history of pulmonary thromboembolism postpartum which was resolved with prolonged anticoagulant therapy. The patient complained for heaviness, tenderness and presence of varicose veins in both legs. The echography revealed bilateral insufficiency of the GSV, with the deep venous system permeable and competent. The AI platform proposed post-thrombotic syndrome as the main diagnosis. CVD was included in the differential diagnosis in all of these three cases.

Discussing the differential diagnosis given by the AI in the patients with venous pathology (Figure 1), the most common was post-thrombotic syndrome, in 93% of the cases, Deep venous thrombosis (87%), Lipedema (81%), Lymphedema (71%), cellulitis (70%), superficial venous thrombosis (35%), compartmental syndrome (32%) and peripheric arteriopathy (32%). Other less frequent diagnosis were arteriovenous malformation (9%), May-Thurner syndrome (5%), pharmacologic edema (1%), lumbar spine stenosis (2%), hypothyroidism (2%), diabetic neuropathy (1%), endometriosis (1%), acrocyanosis (1%), Buerger syndrome (1%). The presence or absence of these diseases in the differential diagnosis depended on the included data in the medical history like sex, age, personal and familiar history, signs and symptoms described in the physical exploration and the findings in the complementary tests. In these patients, CVD was present in the category of "Most likely differential diagnosis" of the AI platform in all of them.

Discussion

This study shows that clinical AI decision support platforms may be a useful and precise tools for the diagnosis of CVD.

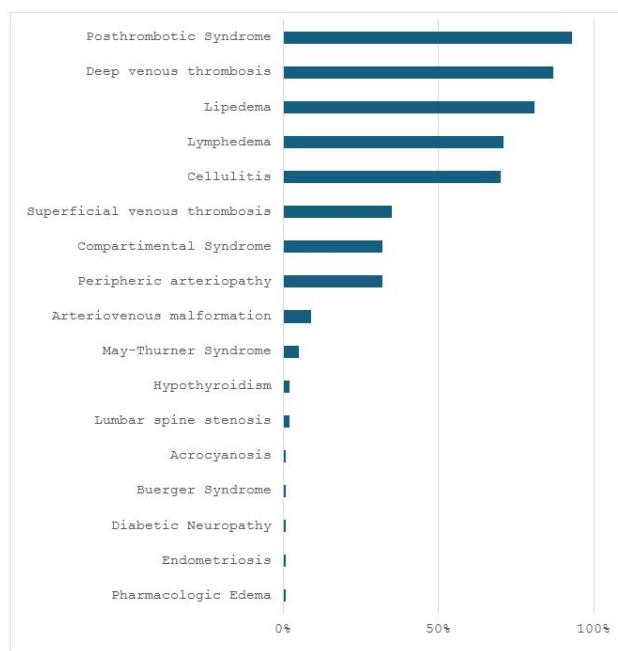


Figure 1. Conditions included in the differential diagnosis with its frequency

Cases in which the AI-generated diagnosis diverged from that of the vascular specialist underscore the importance of detailed and structured data entry to optimize the accuracy and clinical relevance of differential diagnoses. Given that the primary aim of this study was to assess diagnostic concordance, such discrepancies may actually represent instances in which the AI platform produced a more accurate interpretation than the clinician.

These AI platforms are designed to be a complementary tool to help diagnosis and clinical decision making, not substituting the doctor's judgement. They're available from any computer or phone device with internet access. The integration of this tool on the daily work may help to improve differential diagnosis making it more accurate and faster, especially in general services like urgencies or internal medicine. In these settings, the platform could act as a first-line decision support system for general practitioners, assisting with initial diagnostic impressions and guiding the need for specialist referral. This may be particularly beneficial in resource-limited environments or when managing patients with nonspecific or overlapping symptoms.

Medical errors are not infrequent in daily clinical practice, arising not only from misdiagnoses by clinicians but also involving factors related to the patient and the broader healthcare system. The consequences of such errors vary from none to severe with catastrophic consequences for the patient, his family, the healthcare practitioners and cost-effectiveness of the healthcare infrastructure¹⁰. Focusing on the potential errors made by the physicians, the study

of Geoffrey R et al¹¹ explores the cognitive processes of thinking, based in two types: Type 1, the automatic, rapid and intuitive, and Type 2, the slower, logical and analytical thinking. The article also describes where errors can arise and strategies to reduce them. AI platforms healthcare-focused like the one described in this study can play a crucial role to reduce the diagnostic errors in both types of thinking and at the same time, if used correctly, expand the health care's knowledge for future diagnosis.

It's also essential to address the moral and ethical issues that may carry¹², especially related to data protection and responsibility in case of diagnostic errors. More studies are needed to validate and concretize the efficacy and utility of artificial intelligence in the different areas of medicine.

This study has several limitations. First of all, it's a retrospective study, where the diagnosis by the specialist had already been established at the time of the data collection, potentially introducing reference standard bias. The sample is relatively small to take generalized conclusions. Also the majority of the patients had venous pathology. While the AI had a great diagnostic accuracy in CVD, it didn't have it in the other pathologies. It's important to note that while the diagnosis of peripheral arteriopathy was consistent between the specialist and the AI, the kappa coefficient did not reflect a high level of correlation due to the limited representation of this condition in the study sample. The AI's performance is highly dependent on the quality and level of detail of the input data, making it susceptible to operator bias, since all information was manually entered by the investigators. Another important limitation is the lack of external validation. Future research should assess the model's performance on independent datasets from other clinical settings to confirm its generalizability.

The diagnosis of CVD is, in most cases, relatively straightforward for an experienced clinician. It typically relies on well-defined clinical signs and standardized instrumental assessment such as color duplex ultrasound. As a result, the use of CVD as a diagnostic model may not be ideal to thoroughly test the advanced diagnostic capabilities of artificial intelligence. Future studies should aim to extend the analysis to more complex clinical scenarios or conditions with greater symptom overlap, where AI support could truly make a difference. A more robust test of AI capabilities could involve the analysis of duplex ultrasound findings combined with functional data from plethysmography, assessment of edema, classification of venous shunts following the CHIVA protocol, and evaluation of the cost-benefit ratio of various treatment strategies.

There are no conflicts of interest to declare regarding this study. The research was conducted independently, and

no external funding, sponsorships, or affiliations influenced the study design, data collection, analysis, or conclusions.

Conclusion

This study reveals the efficacy of artificial intelligence platforms for the diagnosis of chronic venous disease from clinical human data. With an accuracy of 97% respect to vascular surgeons, its utility to support clinical decision-making is evident. When using this tool it's very important to give detailed and precise data of patient's history to improve its utility to offer a correct differential diagnosis.

More studies are needed to validate and improve the efficacy of artificial intelligence in the different areas of Medicine. The adoption of these technologies may bring a more personalized medicine, accurate and efficient, benefiting both patients and health care professionals.

References

- 1) Krittanawong C, Zhang HJ, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. *J Am Coll Cardiol*. 2017;69(21):2657-2664.
- 2) Thrall JH, Li X, Li Q, et al. Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success. *J Am Coll Radiol*. 2018;15(3):504-508.
- 3) Itchhaporia D. Artificial intelligence in cardiology. *Trends Cardiovasc Med*. 2022;32(1):34-41.
- 4) Hashimoto DA, Witkowski E, Gao L, Meireles O, Rosman G. Artificial intelligence in anesthesiology: Current techniques, clinical applications, and limitations. *Anesthesiology*. 2020;(2):379-394.
- 5) Dossabhoy SS, Ho VT, Ross EG, Rodriguez F, Arya S. Artificial intelligence in clinical workflow processes in vascular surgery and beyond. *Semin Vasc Surg*. 2023;36(3):401-412.
- 6) Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. *Healthc Manag Forum*. 2020;33(1):10-18.
- 7) Glass Health. Glass Health website. <https://glass.health/>. Accessed March 21, 2024.
- 8) Davies AH. The Seriousness of Chronic Venous Disease: A Review of Real-World Evidence. *Adv Ther*. 2019;36(s1):5-12.
- 9) Eklof B, Perrin M, Delis KT, Rutherford RB, Gloviczki P. Updated terminology of chronic venous disorders: The VEIN-TERM transatlantic interdisciplinary consensus document. *J Vasc Surg*. 2009;49(2):498-501.
- 10) Linda T Kohn, Janet M Corrigan and Molla S Donaldson. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academies Press; 2000.
- 11) Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The Causes of Errors in Clinical Reasoning: Cognitive Biases, Knowledge Deficits, and Dual Process Thinking. *Acad Med*. 2017;92(1):23-30.
- 12) Busch F, Adams LC, Bressem KK. Biomedical Ethical Aspects Towards the Implementation of Artificial Intelligence in Medical Education. *Med Sci Educ*. 2023;33(4):1007-1012.

