Identifying Patients With Peripheral Artery Disease Using the Electronic Health Record:

A Pragmatic Approach

Mark Sonderman, MD[a], **Aaron W. Aday, MD, MSc**[b], **Eric Farber-Eger, BS**[c], **Quan Mai, MS**[d], **Matthew S. Freiberg, MD, MS**[e], **David M. Liebovitz, MD**[d], **Philip Greenland, MD**[d], **Mary M. McDermott, MD**[d], **Joshua A. Beckman, MD**[f], **Quinn Wells, MD, PharmD**[e]

[a]Division of Cardiology, Department of Medicine, University of Washington Medical Center, Seattle, Washington, USA;

[b]Division of Cardiovascular Medicine, Vanderbilt Translational and Clinical Cardiovascular Research Center, Vanderbilt University Medical Center, Nashville, Tennessee, USA;

[c]Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, Tennessee, USA;

[d]Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA;

[e]Division of Cardiovascular Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA;

[f]Division of Cardiology, Department of Medicine, University of Texas Southwestern Medical Center, Dallas, Texas, USA.

Abstract

BACKGROUND—Peripheral artery disease (PAD) is underdiagnosed due to poor patient and clinician awareness. Despite this, no widely accepted PAD screening is recommended.

OBJECTIVES—The authors used machine learning to develop an automated risk stratification tool for identifying patients with a high likelihood of PAD.

METHODS—Using data from the electronic health record (EHR), ankle-brachial indices (ABIs) were extracted for 3,298 patients. In addition to ABI, we extracted 60 other patient characteristics and used a random forest model to rank the features by association with ABI. The model identified several features independently correlated with PAD. We then built a logistic regression model to predict PAD status on a validation set of patients (n = 1,089), an external cohort of patients (n =

ADDRESS FOR CORRESPONDENCE: Dr Mark Sonderman, University of Washington Medical Center, 1959 NE Pacific Street, Box 356422, Seattle, Washington, 98195-6422. mark.sonderman@vumc.org.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

APPENDIX For supplemental tables, please see the online version of this paper.

2,922), and a national database (n = 2,488). The model was compared to an age-based and random forest model.

RESULTS—The model had an area under the curve (AUC) of 0.68 in the validation set. When evaluated on an external population using EHR data, it performed similarly with an AUC of 0.68. When evaluated on a national database, it had an AUC of 0.72. The model outperformed an age-based model (AUC: 0.62; $P < 0.001$). A random forest model with inclusion of all 60 features did not perform significantly better (AUC: 0.71; $P = 0.31$).

CONCLUSIONS—Statistical techniques can be used to build models which identify individuals at high risk for PAD using information accessible from the EHR. Models such as this may allow large health care systems to efficiently identify patients that would benefit from aggressive preventive strategies or targeted-ABI screening.

Keywords

ABI; linear models; machine learning; prediction; risk assessment

---

Peripheral artery disease (PAD) affects over 230 million people world-wide.[1,2] PAD reduces walking ability and increases risk of adverse limb events, such as amputation.[2] PAD is also associated with increased risk of adverse cardiovascular events. Despite its high burden, it remains underdiagnosed as most patients do not manifest typical symptoms. Classic intermittent claudication is only found in 10% to 30% of patients and nearly half of patients are atypically symptomatic or report being asymptomatic.[2–4] Diagnosis of PAD has been shown to increase rates of initiation of PAD specific therapies and better outcomes.[5]

Despite this, there is no consensus on PAD screening. The American College of Cardiology and American Heart Association recommend screening in asymptomatic patients who are aged 65 years and older or who are aged 50 to 64 years with risk factors for atherosclerosis with an ankle-brachial index (ABI).[6] The U.S. Preventive Services Task Force (USPSTF) concluded that there was insufficient evidence to recommend screening for PAD in asymptomatic adults, ie, persons aged >18 years.[7] Insurance coverage for preventative health screening without cost sharing to the patient in the United States is tied to USPSTF A and B recommendations.[8] Thus, there is no current mechanism to systemically identify patients at risk for PAD, such as those who are older age, have diabetes, or a history of smoking. In the absence of a population screening strategy, an automated electronic health record (EHR)-based method to identify a group of patients at highest risk of PAD would foster earlier institution of medical therapies, promote pedal examination, and enable investigation into whether EHR-based screening of patients at high risk of PAD improves outcomes.

We used machine learning and logistic regression to develop models to estimate the presence of PAD, as defined by ABI <0.90, in patients without prior ABI testing in their electronic records.[9] These models were validated on internal validation and external test cohorts and their performance was compared to age-based screening. Our goal was to develop a pragmatic model to efficiently identify patients at high risk for PAD based on information entirely contained within the electronic medical record that could be used to facilitate highly targeted screening or initiation of preventive strategies in an undiagnosed population.

METHODS

COHORT FOR TOOL DEVELOPMENT AND VALIDATION.

This research was approved by the Vanderbilt University Institutional Review Board and Northwestern University Institutional Review Board. We collected data on all patients with resting ABI testing present in Vanderbilt University Medical Center's (VUMC) deidentified EHR database, the Synthetic Derivative. There were 75,803 ABI tests for 8,093 patients within the database that were extracted using regular expression. The earliest recorded ABI was chosen for each patient. If multiple tests were present on the earliest date, the lowest single ABI value was chosen. We then performed our analysis on all patients over the age of 18 years (N = 3,298) who had their first ABI measured between January 2015 and October 2020. External model testing was performed on all adult patients (N = 2,922) who had their first ABI measured between January 2015 and December 2020 at Northwestern Medicine (NW). There were no exclusion criteria for these populations. External model testing was also performed on all patients with ABI testing in the 2003 to 2004 National Health and Nutrition Examination Survey (NHANES) database (n = 2,488).[10] All patients with ABI testing in the database were over the age of 40 years and patients were excluded if they weighed over 400 pounds or had bilateral lower extremity amputations.

OUTCOME.

PAD was defined as an ABI <0.90 in either limb on the date of first ABI measurement.

MODEL COVARIATES.

We identified 60 features that were readily extractable from the EHR and had been either identified in prior published data or were biologically plausible as influencers of PAD. We only used unstructured data. These features consisted of patient demographics, vital signs, lab values, diagnoses, and medications. The features were limited to those that were present within the EHR on or before the date of ABI measurement. Related features were combined into aggregate variables (additional detail in Supplemental Tables 1 to 3). Medications were limited to those that were present at any time on the patient's medication list in the 365 days prior to ABI measurement. There was no limit to the look back period for other features. Missing data were imputed with the median of the value across the population.

FEATURE PRIORITIZATION AND MODEL CREATION.

We used the 60 vascular disease associated features to train a random forest algorithm for predicting patient ABIs. The random forest assigned a variable importance to each feature which described the strength of that feature's independent association with ABI (Supplemental Table 1). Vascular medicine specialists (A.W.A., J.A.B.) were consulted to build 5 models using features with high variable importance in the random forest algorithm and completeness within the EHR while excluding overlapping features (Supplemental Table 4). An age-based model was built for comparison in which a test statistic ranging from 0 to 1 was linearly proportional to the patient's age.

MODEL IMPLEMENTATION, EVALUATION, AND VALIDATION.

The models described above were trained on a cohort of patients from VUMC (n = 2,209) to build logistic regression equations that described the chosen feature's impact on a patient's risk of having an ABI <0.90. After training, the logistic regression models were then used to estimate presence of PAD on 3 separate cohorts of patients: an internal VUMC cohort (n = 1,089), an external cohort from NW, another academic medical center (n = 2,922), and an external cohort from the 2003 to 2004 NHANES database, a nationally-representative survey designed to assess the health and nutritional status of adults and children in the United States.[11]

Each model outputs a continuous test statistic for each patient ranging from 0 to 1. A higher test statistic indicates a higher probability of a patient having an ABI <0.90. The test statistic threshold for identifying patients as high risk of PAD can be customized to the population being studied to prioritize high positive predictive value (PPV) vs negative predictive value (NPV). A patient will be identified as high risk of having an abnormally low ABI if their test statistic is greater than the threshold chosen.

STATISTICAL ANALYSIS.

All statistical analyses were performed using R (version 3.6.3).[12] The random forest used conditional interference trees as base learners using the partykit package.[13] Conditional interference trees were used because variable importance in traditional random forests is biased in favor of features that contain more levels (ie, continuous variables such as lab values or vital signs).[13,14] Conditional interference trees provide unbiased variable selection when there are both continuous and categorical features. Logistic regression models were built for validation as random forests require more computational time that make implementation within the EHR unrealistic. Additionally, random forests are black box models that are less interpretable for individual clinicians due to the absence of coefficients describing the magnitude of a feature's effect on the outcome and increased interaction between features as the size of the model grows.[15,16]

Receiver operating characteristic (ROC) curves were built using predictions from the logistic regression and area under the curve (AUC) calculations were performed using the pROC package.[17] All additional statistical analyses including model-level sensitivity, specificity, PPV, and NPV were calculated for model test statistic thresholds at percentiles ranging from 0 to 1.

RESULTS

STUDY POPULATION.

At VUMC, we identified 3,298 unique patients with a first ABI measured between January 2015 and October 2020, of which 1,545 (47%) <0.90. Patient characteristics are presented in Table 1. Among the study population, 76% were White, 17% were Black, 57% were male, the mean age was 57.8 ± 16.6 years, and the mean body mass index (BMI) was 29 ± 6.6 kg/m^2. Comorbidities were common, 53% of patients had hypertension, 20% had diabetes mellitus, 26% had a history of smoking, and 30% had a prior diagnosis of PAD.

MODEL PREDICTION RESULTS.

The 5 different logistic regression models performed similarly based on the ROC curve with AUCs ranging from 0.67 to 0.69 (Supplemental Table 4). Given similar performance across all 5 models, the objective model (AUC: 0.68) was selected because of its parsimony and simplicity and will form the basis of this report (Central Illustration). The model consisted of the following features: age, gender, systolic blood pressure, diastolic blood pressure, pulse pressure, antidiabetic medication use, antihypertensive medication use, history of smoking, and total cholesterol to high-density lipoprotein cholesterol ratio. An exclusively age-based screening model performed significantly worse in this population as compared to the objective model (AUC: 0.62; $P < 0.001$). A model built using a random forest with inclusion of all features did not perform significantly better (AUC: 0.71; $P = 0.31$). The model performance was not significantly different between White (AUC: 0.66) and non-White patients (AUC: 0.72; $P = 0.10$). The model performed significantly better with male patients (AUC: 0.71) than female patients (AUC: 0.64; $P = 0.03$).

The test characteristic was correlated with a patient's likelihood of having PAD, and this relationship was observed for both low- and high-risk patients (Figure 1). At a test statistic threshold of 0.3, the model had a sensitivity of 90%, specificity of 26%, PPV of 52%, NPV of 74%, and it identified 888 of 1,089 patients as candidates for ABI testing due to high risk for having PAD. At a threshold of 0.5, the model had a sensitivity of 58%, specificity of 69%, PPV 63%, and NPV 65% and identified 482 patients as candidates for ABI testing. At a threshold of 0.7, the model had a sensitivity of 11%, specificity 97%, PPV 76%, and NPV 54% and identified 74 patients for ABI testing.

EXTERNAL TESTING.

The external cohort at Northwestern consisted of 2,922 patients, of which 1,134 (39%) had an ABI <0.90. Characteristics of these patients are presented in Table 1. In this cohort, 80% of the patients were White, 10% were Black, 56% were male, the mean age was 66.9 ± 12.8 years, and the mean BMI was 29.1 ± 6.4 kg/m^2. Hypertension was present in 70%, 36% had diabetes mellitus, and 17% had a history of smoking.

Evaluating the model on the external NW test cohort yielded an AUC of 0.68. To achieve similar test statistics on the external cohort as on the internal cohort required setting a higher threshold to identify a patient as having PAD. Whereas a threshold of 0.5 for internal data yielded a sensitivity of 58% and specificity of 68%, a threshold of 0.53 was required to have a sensitivity of 59% and specificity of 68% on external patients. At a threshold of 0.5, the model would flag 44% of the internal patient population whereas it would flag 50% of the external population. The variance arises from higher rates of traditional PAD risk factors in the external cohort (Table 1). Despite an increased prevalence of risk factors such as diabetes and hypertension, the external cohort had a lower prevalence of ABI <0.90 in their population (39% vs 47%).

This 2003 to 2004 NHANES database consisted of 2,488 patients, of which 227 (9%) had an ABI <0.90. Characteristics of these patients are presented in Table 1. In this cohort, 58% of patients were White, 18% were Black, 51% were male, the mean age was 61.1 + 13.2

years, and the mean BMI was 28. 5 ± 5.5 kg/m^2. Hypertension was present in 39%, 14% had diabetes mellitus, and 55% had a history of smoking.

Evaluating the model on the NHANES database yielded an AUC of 0.72. To achieve similar test statistics on the NHANES cohort required setting a higher threshold to identify a patient as having PAD. A threshold of 0.55 yields a sensitivity of 60% and specificity of 71% on NHANES patients, similar to results for the internal data at a threshold of 0.5 and external data at 0.53. At a threshold of 0.5, the model would flag 42% of the NHANES population for ABI screening (as compared to 44% of the internal population and 50% of the external population). If more selective screening was desired in this population with a lower prevalence of PAD, the threshold could be increased to 0.68 and would flag 10% of the population for ABI screening.

DISCUSSION

The aim of our project was to demonstrate the feasibility of using both machine learning (random forest) and logistic regression to develop a risk stratification tool for identifying patients at increased risk for PAD suitable for ABI testing. Confirmation bias within the population was minimized by using a patient's first ABI within the EHR and excluding variables from inclusion in the model that may reflect a diagnosis of PAD (vascular disease diagnoses, prior vascular interventions, or cilostazol use). The model estimated abnormally low ABI in 2 separate academic medical centers with an AUC of 0.68. The model performed similarly when validated in the NHANES database with an AUC of 0.72. The model outperformed an age-based ABI screening approach. Despite the model's consistent performance between medical centers, the different thresholds needed to achieve similar performance between centers show that the ideal test statistic threshold for identifying high-risk patients varies depending on a population's risk factors, regional PAD prevalence, and desired rate of discovery.

Prior studies have also used machine learning for PAD prediction. Ross et al[18] developed a random forest algorithm which used 120 patient characteristics that had an AUC of 0.87 for PAD prediction in 1,047 patients not previously diagnosed with PAD presenting for coronary angiography. Our work advances screening importantly. First, 40% of all PAD patients do not have evidence of other atherosclerosis.[19] Second, the presence of atherosclerosis in one vascular bed increases the risk significantly of its presence in another and incidence of adverse events.[20–24] The high accuracy of this study certainly demonstrates the power of using machine learning to predict PAD status using a strategy of very high-risk patients (76% of patients in this study had coronary artery disease), but we believe the power of the EHR is to extend these methods to all patients within a medical center who have clinical data within the EHR.

Therefore, it is possible to build machine learning models with improved accuracy over our models described in this paper by limiting the population to higher risk cohorts or using time-intensive, manually collected data rather than extracting it from the EHR. But these methods are difficult to implement outside of the research setting and limit the population in who PAD can be diagnosed prior to adverse events. Unstructured data from

the EHR can be analyzed retrospectively and built into a model, but it is less practical for prospective implementation because of the variation in identification and prevalence across EHR platforms. Our model is capable of risk stratifying patients even when components of the model are missing. We sacrifice predictive accuracy modestly in exchange for a greater discovery of cases. We believe this strikes the right balance between accuracy and utility. The risks associated with testing are near nil while the benefits of diagnosis are considerable. Fewer features, then, foster easier implementation for routine utilization in clinical practice. Including fewer features in our model which reduced overfitting on internal data is demonstrated by its consistent performance when evaluated on external data which is not typically seen in clinical prediction models.[25] Over 40% of clinical prediction models described in the literature are not externally evaluated and when they are, the AUC typically decreases by 11%.[26]

The benefit of the model is its continuous test statistic which captures a patient's individualized risk rather than giving a binary response (screen or do not screen) as with the American College of Cardiology/American Heart Association screening guidelines or evaluated in the DANCAVAS (Danish Cardiovascular Screening Trial).[27] This allows the model to be tuned to change the number of patients identified through screening and maximize true discovery while minimizing missing patients and false positive tests. We would prioritize high NPV when implementing the model for screening, and therefore, pick a low test statistic threshold. Our model discovered 302 people with PAD for every 481 people identified for screening at a model threshold of 0.5 in a population of 1,089 patients with a prevalence of PAD of 47%; 215 patients with PAD were not identified for screening (Table 2). This was significantly better than age-based ABI screening and had similar performance across all models. In a population that underwent ABI testing without regard to symptoms (NHANES), our model discovered 167 people with PAD for every 1,055 identified for screening in a population of 2,488 patients with a PAD prevalence of 9%; 60 patients with PAD were not identified for screening. The benefit of the model in this paper is its simplicity and pragmatic use case.

Future directions include addressing model performance while maintaining our goal of use simplicity. The inclusion of diagnostic data (imaging, electrocardiograms, and physiologic data) may improve performance by finding a higher-risk cohort.[28] Adding data of similar type as already included in this analysis would unlikely improve performance and would more likely lead to overfitting data with reduced external validity.[16] Our finding that model-specific adjustments (comparing logistic regression vs machine learning models) did not significantly improve model performance suggests that improving data quality within the model is higher yield than building more advanced machine learning models, a finding consistent with previous research.[29] This paper serves as a proof of concept for building data driven models using only information contained within the EHR. Prospective implementation of our model with evaluation of both its performance and effect on outcomes is the most important next step to determine its clinical utility.

Author Manuscript

STUDY LIMITATIONS.

First, the results reported here cannot be generalized to an asymptomatic population. Second, the analyses did not discriminate by ordering indication. Third, the model aimed to use first measured ABIs to capture index diagnosis of PAD but 30% of the internal population had previously been diagnosed with PAD. Prospective implementation of this model would address each of these limitations by evaluating its performance in an asymptomatic population without a prior PAD diagnosis. Lastly, the population used to develop this model was representative of the national Black population, but the non-Hispanic, White population was overrepresented as compared to the local and national population.[30] These variances may impact model accuracy in Hispanic and Asian populations.[28]

CONCLUSIONS

PAD is a disease associated with high morbidity and mortality but is underdiagnosed as classic symptoms of claudication are rarely present and screening opportunities are not covered. This paper describes the development of a model built using both machine learning and logistic regression techniques that identified patients at high risk of having PAD that was validated retrospectively at 2 academic medical centers. Further studies to prospectively validate the model are needed to evaluate its efficacy in identifying PAD in asymptomatic populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

FUNDING SUPPORT AND AUTHOR DISCLOSURES

ABBREVIATIONS AND ACRONYMS

ABI	ankle-brachial index
AUC	area under the curve
BMI	body mass index
EHR	electronic health record

NPV	negative predictive value
NW	Northwestern Medicine
PAD	peripheral artery disease
PPV	positive predictive value
ROC	receiver operating characteristic

REFERENCES

1. Song P, Rudan D, Zhu Y, et al. Global, regional, and national prevalence and risk factors for peripheral artery disease in 2015: an updated systematic review and analysis. Lancet Glob Health. 2019;7(8):e1020–e1030. 10.1016/S2214-109X(19)30255-4 [PubMed: 31303293]

2. Polonsky TS, McDermott MM. Lower extremity peripheral artery disease without chronic limb-threatening ischemia: a review. JAMA. 2021;325(21):2188. 10.1001/jama.2021.2126 [PubMed: 34061140]

3. McDermott MM. Lower extremity manifestations of peripheral artery disease: the pathophysiologic and functional implications of leg ischemia. Circ Res. 2015;116(9):1540–1550. 10.1161/CIRCRESAHA.114.303517 [PubMed: 25908727]

4. Criqui MH, Matsushita K, Aboyans V, et al. Lower extremity peripheral artery disease: contemporary epidemiology, management gaps, and future directions: a scientific statement from the American Heart Association. Circulation. 2021;144(9):e171–e191. 10.1161/CIR.0000000000001005 [PubMed: 34315230]

5. Lindholt JS, Søgaard R. Population screening and intervention for vascular disease in Danish men (VIVA): a randomised controlled trial. Lancet. 2017;390(10109):2256–2265. 10.1016/S0140-6736(17)32250-X [PubMed: 28859943]

6. Gerhard-Herman MD, Gornik HL, Barrett C, et al. 2016 AHA/ACC guideline on the management of patients with lower extremity peripheral artery disease: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2017;135(12):e686–e725. 10.1161/CIR.0000000000000470 [PubMed: 27840332]

7. US Preventive Services Task Force, Curry SJ, Krist AH, et al. Screening for peripheral artery disease and cardiovascular disease risk assessment with the ankle-brachial index: US Preventive Services Task Force recommendation statement. JAMA. 2018;320(2):177. 10.1001/jama.2018.8357 [PubMed: 29998344]

8. Seiler N, Malcarney MB, Horton K, Dafflitto S. Coverage of clinical preventive services under the affordable care act: from law to access. Public Health Rep. 2014;129(6):526–532. 10.1177/003335491412900611 [PubMed: 25364055]

9. Aboyans V, Criqui MH, Abraham P, et al. Measurement and interpretation of the ankle-brachial index: a scientific statement from the American Heart Association. Circulation. 2012;126(24):2890–2909. 10.1161/CIR.0b013e318276fbcb [PubMed: 23159553]

10. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data. Hyattsville, MD. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2003–2004.

11. Marwick C. How do they conduct "N-HANES," anyway? JAMA. 1990;263(19):2581.

12. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc; 2019.

13. Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. BMC Bioinf. 2008;9(1):307. 10.1186/1471-2105-9-307

14. Hothorn T. Survival ensembles. Biostatistics. 2005;7(3):355–373. 10.1093/biostatistics/kxj011 [PubMed: 16344280]

15. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019;28(3):231–237. 10.1136/bmjqs-2018-008370

16. Deo RC. Machine learning in medicine. Circulation. 2015;132(20):1920–1930. 10.1161/CIRCULATIONAHA.115.001593 [PubMed: 26572668]

17. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinf. 2011;12(1):77. 10.1186/1471-2105-12-77

18. Ross EG, Shah NH, Dalman RL, Nead KT, Cooke JP, Leeper NJ. The use of machine learning for the identification of peripheral artery disease and future mortality risk. J Vasc Surg. 2016;64(5): 1515–1522.e3. 10.1016/j.jvs.2016.04.026 [PubMed: 27266594]

19. Hirsch AT, Criqui MH, Treat-Jacobson D, et al. Peripheral arterial disease detection, awareness, and treatment in primary care. JAMA. 2001;286(11):1317–1324. 10.1001/jama.286.11.1317 [PubMed: 11560536]

20. Colantonio LD, Hubbard D, Monda KL, et al. Atherosclerotic risk and statin use among patients with peripheral artery disease. J Am Coll Cardiol. 2020;76(3):251–264. 10.1016/j.jacc.2020.05.048 [PubMed: 32674789]

21. Gutierrez JA, Mulder H, Jones WS, et al. Polyvascular disease and risk of major adverse cardiovascular events in peripheral artery disease: a secondary analysis of the EUCLID trial. JAMA Netw Open. 2018;1(7):e185239. 10.1001/jamanetworkopen.2018.5239 [PubMed: 30646395]

22. Gutierrez JA, Aday AW, Patel MR, Jones WS. Polyvascular disease: reappraisal of the current clinical landscape. Circ Cardiovasc Interv. 2019;12(12):e007385. 10.1161/CIRCINTERVENTIONS.119.007385 [PubMed: 31833412]

23. Subherwal S, Bhatt DL, Li S, et al. Polyvascular disease and long-term cardiovascular outcomes in older patients with non-ST-segment-elevation myocardial infarction. Circ Cardiovasc Qual Outcomes. 2012;5(4):541–549. 10.1161/CIRCOUTCOMES.111.964379 [PubMed: 22715460]

24. Bonaca MP, Gutierrez JA, Cannon C, et al. Polyvascular disease, type 2 diabetes, and long-term vascular risk: a secondary analysis of the IMPROVE-IT trial. Lancet Diabetes Endocrinol. 2018;6(12):934–943. 10.1016/S2213-8587(18)30290-0 [PubMed: 30396865]

25. Kernbach JM, Staartjes VE. Machine learning-based clinical prediction modeling – A practical guide for clinicians. arXiv. 2020. 10.48550/ARXIV.2006.15069

26. Wessler BS, Nelson J, Park JG, et al. External validations of cardiovascular clinical prediction models: a large-scale review of the literature. Circ Cardiovasc Qual Outcomes. 2021;14(8): e007858. 10.1161/CIRCOUT-COMES.121.007858 [PubMed: 34340529]

27. Lindholt JS, Søgaard R, Rasmussen LM, et al. Five-year outcomes of the Danish cardiovascular screening (DANCAVAS) trial. N Engl J Med. 2022;387(15):1385–1394. 10.1056/NEJMoa2208681 [PubMed: 36027560]

28. Manlhiot C, van den Eynde J, Kutty S, Ross HJ. A primer on the present state and future prospects for machine learning and artificial intelligence applications in cardiology. Can J Cardiol. 2022;38(2):169–184. 10.1016/j.cjca.2021.11.009 [PubMed: 34838700]

29. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019;110:12–22. 10.1016/j.jclinepi.2019.02.004 [PubMed: 30763612]

30. U.S. Census Bureau. QuickFacts: United States. 2021. Accessed August 10, 2022. http://www.census.gov/quickfacts/fact/table/US/PST045221

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE:

Ankle-brachial index testing is indicated in patients with exertional leg symptoms and risk factors for peripheral artery disease.

COMPETENCY IN PROFESSIONALISM:

When evaluating a clinical prediction model, the external validity of the model is dependent on whether the diversity of the population used to train the model reflects the population in which it will be used.

COMPETENCY IN SYSTEMS-BASED PRACTICE:

Preventive screening that has an "A" or "B" recommendation from the U.S. Preventive Services Task Force must be covered by private insurance without cost sharing to the patient.

TRANSLATIONAL OUTLOOK:

Evaluation of clinical prediction models prospectively is needed to evaluate their efficacy as a screening tool.
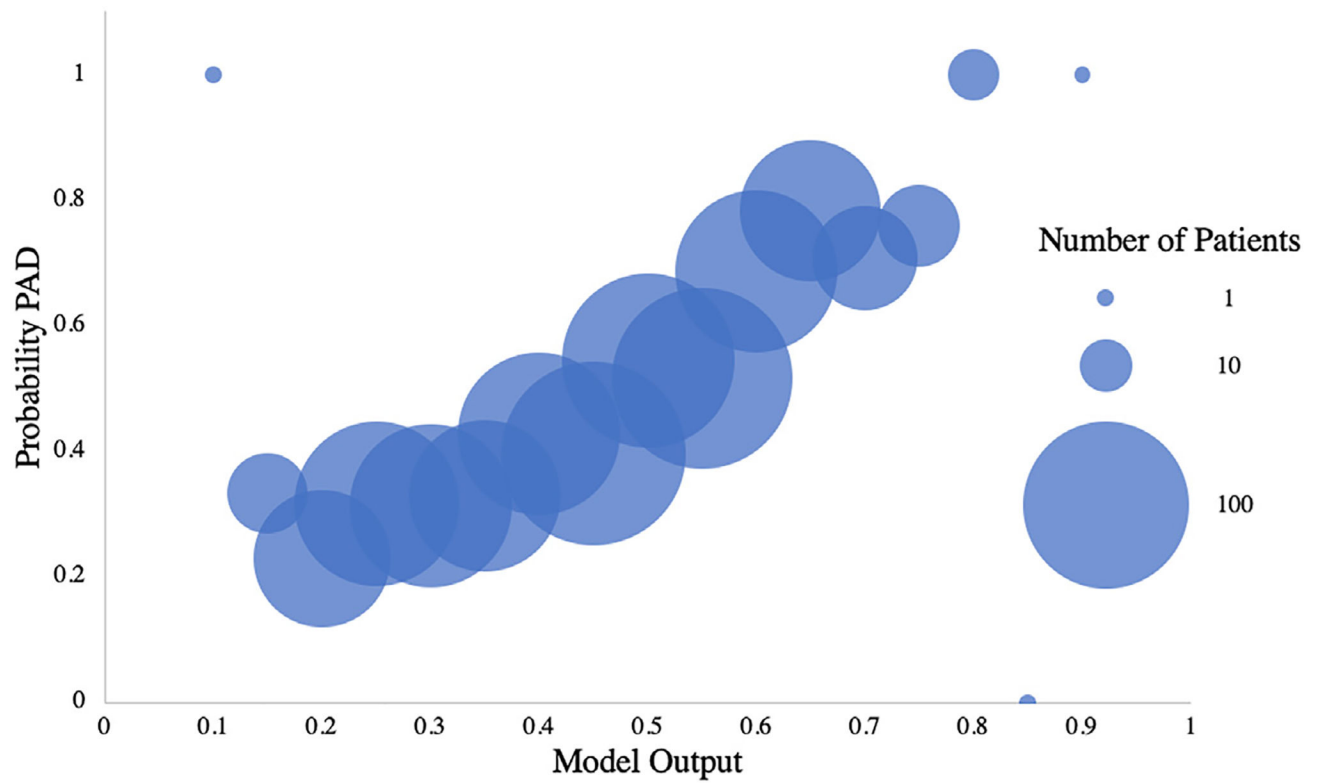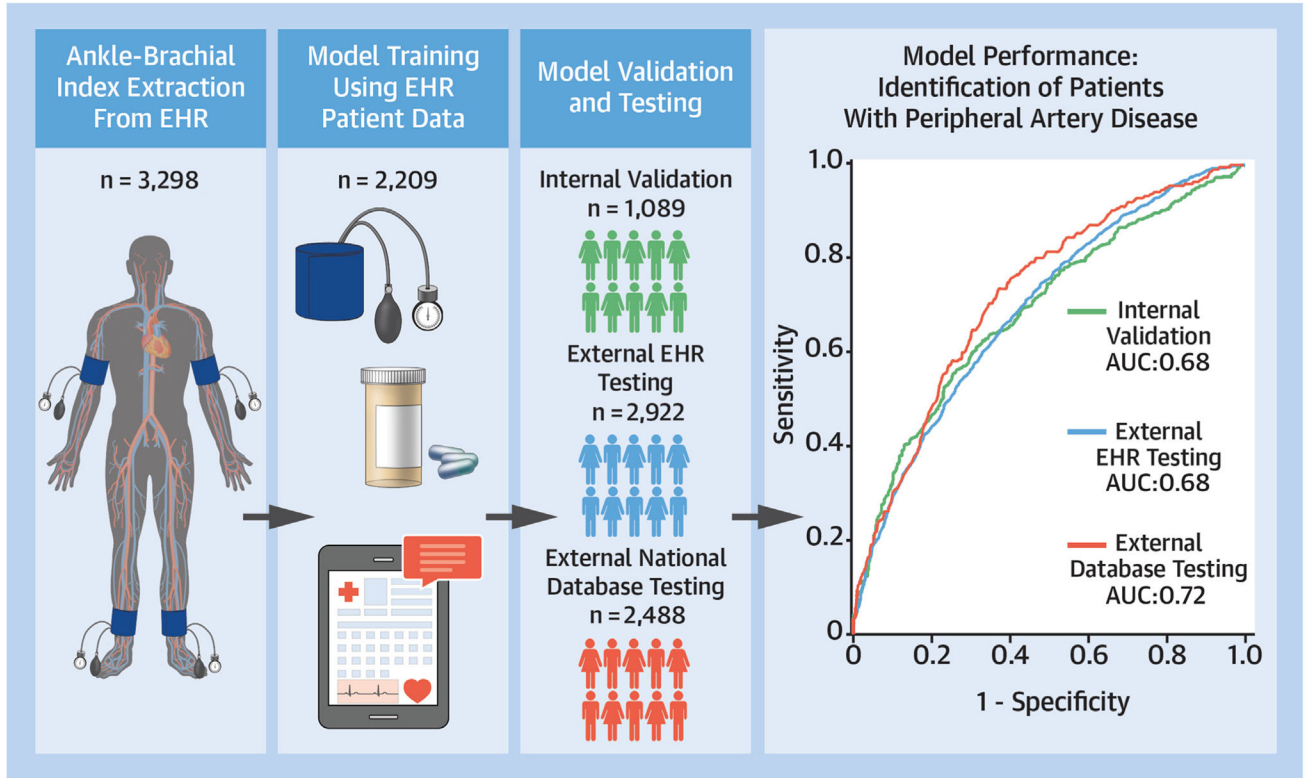
FIGURE 1.

Model Output Vs ABI Validated PAD Status

Model calibration demonstrated by mean test statistic as compared to the percentage of patients within each quantile that had abnormally low ABIs. Marker size is proportional to number of patients with each quantile. ABI = ankle-brachial index; PAD = peripheral artery disease.

CENTRAL ILLUSTRATION.

Peripheral Artery Disease Predictor Model Development

Ankle-brachial indices (ABIs) were extracted from the EHR using regular expression for 3,298 patients. A random forest was then used to analyze 60 patient-specific features (demographics, vital signs, laboratory values, diagnoses, and medication use) to identify features most strongly associated with low ABI. These features were used with a training cohort of patients to build 5 logistic regression models. A single model was chosen and evaluated on both internal validation and external test cohorts (EHR data and National Health and Nutrition Examination Survey) that were not used for model training to evaluate how the models approximated PAD risk. Model performance was evaluated using receiver operating characteristic curves. EHR = electronic health record.

TABLE 1

Baseline Characteristics of Patients at Both Clinical Sites and Within the NHANES Database

	Internal (n = 3,298)	External (n = 2,922)	NHANES (n = 2,488)
ABI	0.81 ± 0.33	0.91 ± 0.29	1.07 ± 0.15
Age (y)	57.8 ± 16.6	66.9 ± 12.8	61.1 ± 13.2
Male	1,878 ± 57%	1,650 ± 56%	1,268 ± 51%
Race/ethnicity			
Caucasian	2,423 (73%)	2,215 (76%)	1,436 (58%)
Black	549 (17%)	303 (10%)	454 (18%)
Asian	19 (1%)	54 (2%)	Not reported
Hispanic	80 (2%)	125 (4%)	534 (21%)
Other/missing	227 (7%)	225 (8%)	64 (3%)
BMI	29.2 ± 6.6	29.1 ± 6.4	28.5 ± 5.5
Ever smoker (% yes)	858 (26%)	489 (17%)	1,358 (55%)
SBP (mm Hg)	129 ± 21	132 ± 19	133 ± 21
DBP (mm Hg)	72 ± 14	72 ± 11	72 ± 11
DM (% yes)	662 (20%)	1,061 (36%)	360 (14%)
HTN (% yes)	1,764 (53%)	2,036 (70%)	974 (39%)
GFR	79 ± 33	62 ± 21	76 ± 20

Values are mean ± SD or n (%).

ABI = ankle-brachial index; BMI = body mass index; DBP = diastolic blood pressure; DM = diabetes mellitus diagnosis; GFR = glomerular filtration rate; HTN = hypertension diagnosis; NHANES = National Health and Nutrition Examination Survey; SBP = systolic blood pressure.

TABLE 2

Model Performance vs Age-Based Screening With Different Datasets[a]

Metric	Model Internal (n = 1,089)	Age-Based Internal (n = 1,089)	Model External NW (n = 2,922)	Age-Based External NW (n = 2,922)	Model NHANES (n = 2,488)	Age-Based NHANES (n = 2,488)
Screened[b]	481	472	1,217	1,230	789	740
Diagnosed[c]	302	270	649	550	137	133
Missed[d]	215	247	447	547	90	94
Accuracy	64%	59%	65%	57%	70%	70%

[a]To achieve similar test statistics between models and datasets, a threshold of 0.5 was used for the model with internal data, 0.62 for age-based screening with internal data, 0.53 for the model with external NW data, 0.75 for age-based screening with external NW data, 0.55 for the model with NHANES data, and 0.65 for age-based screening with NHANES data.

[b]Screened: number of patients with a test statistic above screening threshold that would receive ABI testing.

[c]Diagnosed: number of patients screened that would have an ABI <0.90 (true positive).

[d]Missed: number of patients with an ABI <0.90 that would not undergo ABI testing (false negative).

ABI = ankle-brachial index; NHANES = National Health and Nutrition Examination Survey; NW = Northwestern Medicine.