Developing Trustworthy Artificial Intelligence Models to Predict Vascular Disease Progression: the VASCUL-AID-RETRO Study Protocol

Lotte Rijken, MSc[1,2,3] (iD), Sabrina Zwetsloot, MD[1,2] (iD), Stefan Smorenburg, MSc[1,2] (iD),
Jelmer Wolterink, PhD[4] (iD), Ivana Išgum, Phd, Prof[5,6,7], Henk Marquering, PhD, Prof[5,6] (iD),
Jan van Duivenvoorde, Bsc[1,2], Corrette Ploem, PhD, Prof[8], Roosmarie Jessen, MSc[8] (iD),
Fabio Catarinella, MD[9], Regent Lee, MD, PhD, FRCS[10], Katarzyna Bera, MD, PhD[10] (iD),
Jenny Buisan, MSc[10], Ping Zhang, Bsc[10], Marina Dias-Neto, MD, PhD, FEBVS[11,12] (iD),
Juliette Raffort, MD, PhD, HDR[13,14,15], Fabien Lareyre, MD, PhD, HDR[15,16] (iD),
Catelijne Muller, LLM[17], Igor Koncar, MD, PhD[18,19] (iD), Ivan Tomic, MD, PhD[18,19],
Maja Živković, PhD[20], Tamara Djuric, PhD[20] (iD), Aleksandra Stankovic, PhD[20],
Maarit Venermo, MD, PhD, Prof[21,22] (iD), Riikka Tulamo, MD, PhD[21,22],
Christian-Alexander Behrendt, MD, PhD[23] (iD), Noeska Smit, PhD, Prof[24,25] (iD),
Marlies Schijven, MD, PhD, Prof[3,26,27] (iD), Bert-Jan van den Born, MD, PhD, Prof[28,29],
Ronak Delewi, MD, PhD[30,31], Vincent Jongkind, MD, PhD, FEBVS[1,32] (iD),
Venkat Ayyalasomayajula, PhD[1,2] (iD), and Kak Khee Yeung, MD, PhD, FEBVS, Prof[1,2] (iD)

Abstract

Introduction: Abdominal aortic aneurysms (AAAs) and peripheral artery disease (PAD) are two vascular diseases with a significant risk of major adverse cardiovascular events and mortality. A challenge in current disease management is the unpredictable disease progression in individual patients. The VASCUL-AID-RETRO study aims to develop trustworthy multimodal predictive artificial intelligence (AI) models for multiple tasks including risk stratification of disease progression and cardiovascular events in patients with AAA and PAD. **Methods:** The VASCUL-AID-RETRO study will collect data from 5000 AAA and 6000 PAD patients across multiple European centers of the VASCUL-AID consortium using electronic health records from 2015 to 2024. This retrospectively-collected data will be enriched with additional data from existing biobanks and registries. Multimodal data, including clinical records, radiological imaging, proteomics, and genomics, will be collected to develop AI models predicting disease progression and cardiovascular risks. This will be done while integrating the international ethics guidelines and legal standards for trustworthy AI, to ensure a socially-responsible data integration and analysis. **Proposed Analyses:** A consensus-based variable list of clinical parameters and core outcome set for both diseases will be developed through meetings with key opinion leaders. Blood, plasma, and tissue samples from existing biobanks will be analyzed for proteomic and genomic variations. AI models will be trained on segmented AAA and PAD artery geometries for estimation of hemodynamic parameters to quantify disease progression. Initially, risk prediction models will be developed for each modality separately, and subsequently, all data will be combined to be used as input to multimodal prediction models. During all processes, data security, data quality, and ethical guidelines and legal standards will be carefully considered. As a next step, the developed models will be further adjusted with prospective data and internally validated in a prospective cohort (VASCUL-AID-PRO study). **Conclusion**: The VASCUL-AID-RETRO study will utilize advanced AI techniques and integrate clinical, imaging, and multi-omics data to predict AAA and PAD progression and cardiovascular events.

Clinical Trial Registration: The VASCUL-AID-RETRO study is registered at www.clinicaltrials.gov under the identification number NCT06206369.

Clinical Impact
The VASCUL-AID-RETRO study aims to improve clinical practice of vascular surgery by developing artificial intelligence-driven multimodal predictive models for patients with abdominal aortic aneurysms or peripheral artery disease, enhancing personalized medicine. By integrating comprehensive data sets including clinical, imaging, and multi-omics data, these models have the potential to provide accurate risk stratification for disease progression and cardiovascular events. An innovation lies in the extensive European data set in combination with multimodal analyses approaches, which enables the development of advanced models to facilitate better understanding of disease mechanisms and progression. For clinicians, this means that more precise, individualized treatment plans can be established, ultimately aiming to improve patient outcomes.

[1]Department of Surgery, Amsterdam University Medical Center, Location Vrije Universiteit, Amsterdam, The Netherlands
[2]Atherosclerosis and Ischemic Syndromes, Amsterdam Cardiovascular Sciences, Amsterdam University Medical Center, Amsterdam, The Netherlands
[3]Digital Health Amsterdam Public Health, Amsterdam University Medical Center, Amsterdam, The Netherlands
[4]Department of Applied Mathematics, Technical Medical Centre, University of Twente, Enschede, The Netherlands
[5]Department of Biomedical Engineering and Physics, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[6]Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[7]Informatics Institute, Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands
[8]Department of Ethics, Law and Humanities, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[9]Brightfish BV, Hoofddorp, The Netherlands
[10]Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
[11]Department of Angiology and Vascular Surgery, Centro Hospitalar Universitário de São João, Porto, Portugal
[12]UnIC@RISE, Department of Surgery and Physiology, Faculty of Medicine, University of Porto, Porto, Portugal
[13]Clinical Chemistry Laboratory, University Hospital of Nice, Nice, France
[14]Institute 3IA Côte d'Azur, Université Côte d'Azur, Nice, France
[15]CNRS, UMR7370, LP2M, Université Côte d'Azur, Nice, France
[16]Department of Vascular Surgery, Hospital of Antibes Juan-les-Pins, Antibes, France
[17]ALLAI, Amsterdam, The Netherlands
[18]Faculty of Medicine, University of Belgrade, Belgrade, Serbia
[19]Clinic for Vascular and Endovascular Surgery, Clinical Center of Serbia, Belgrade, Serbia
[20]Laboratory for Radiobiology and Molecular Genetics, VINCA Institute of Nuclear Sciences—National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia
[21]Department of Vascular Surgery, Helsinki University Hospital, Helsinki, Finland
[22]Department of Vascular Surgery, University of Helsinki, Helsinki, Finland
[23]Department of Vascular and Endovascular Surgery, Asklepios Clinic Wandsbek, Asklepios Medical School, Hamburg, Germany
[24]Department of Informatics, University of Bergen, Bergen, Norway
[25]Department of Radiology, Mohn Medical Imaging and Visualization Centre, Haukeland University Hospital, Bergen, Norway
[26]Department of Surgery, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[27]Amsterdam Gastroenterology Endocrinology Metabolism, Amsterdam University Medical Center, Amsterdam, The Netherlands
[28]Department of Public and Occupational Health, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[29]Department of Vascular Medicine, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[30]Department of Cardiology, Amsterdam University Medical Center, Location University of Amsterdam, Amsterdam, The Netherlands
[31]Amsterdam Cardiovascular Sciences, Amsterdam, The Netherlands
[32]Microcirculation, Amsterdam Cardiovascular Sciences, Amsterdam, The Netherlands

Corresponding Author:
Lotte Rijken, Department of Surgery, Amsterdam University Medical Center, Location Vrije Universiteit, Amsterdam UMC Location VUmc, Meibergdreef 9, Amsterdam, 1105 AZ, The Netherlands.
Email: l.rijken@amsterdamumc.nl

Introduction

Abdominal aortic aneurysms (AAAs) and peripheral artery disease (PAD) are prevalent vascular disorders with a large impact on patients' health.[1,2] AAAs are often asymptomatic until they rupture, which is highly lethal with a mortality of 85% to 90%.[3] Once diagnosed, current treatment guidelines advise surgical intervention based on aneurysm diameter (5.5 cm in men and 5.0 cm in women) or aneurysm growth rate (>10 mm/year).[4,5] However, this "one-size-fits-all" approach fails for many individuals because small aneurysms can still rupture.[6] Besides, AAA patients have a high prevalence of myocardial infarction, stroke, heart failure, and coronary artery diseases.[7,8] PAD is characterized by a spectrum of clinical manifestations, including claudication, impaired health-related quality of life, and functional limitations that can progress to critical limb-threatening ischemia.[9] In addition, PAD is a significant risk factor for major adverse cardiovascular events (MACE) such as myocardial infarction, stroke or cardiovascular death, and the need for revascularization procedures or amputation, ultimately increasing morbidity and mortality risks.[9,10]

One of the main challenges in current management of these two vascular diseases is the inter-patient uncertainty in disease progression, leading up to events such as aneurysm rupture in AAA patients, or the need for artery revascularization or amputation in PAD patients. Moreover, the proportion of patients with a high risk of MACE is estimated to be 25%[8] and 23%[11] in AAA and PAD, respectively. Secondary preventive measures for key modifiable risk factors (eg, diabetes, hypertension, and smoking) have proven their efficacy, but they remain to be adequately implemented in clinical practice.[12,13] Previous research has indicated that implementing optimal cardiovascular risk management (CVRM) strategies could effectively reduce the 10-year risk of MACE in AAA patients from 43% to 14%[12]. Nonetheless, the tools to predict patient-specific progression of these diseases, prognosis, and outcomes of interventions are lacking. Therefore, the development of reliable methods to improve the implementation of evidence-based patient-specific management is paramount and could potentially lead to more commitment of patients and physicians for preventive disease management.

Artificial intelligence (AI) is a transformative avenue that enables more precise quantitative analyses within large patient data sets. AI-powered tools have the potential to improve predicting the risk of AAA growth and rupture, allowing for more targeted treatment approaches.[14] Earlier research has explored the contribution of machine learning algorithms to predict AAA growth using patient data and characteristics such as blood vessel elasticity and aortic diameter with promising results.[15] More recent studies developed AI algorithms that analyze patient clinical and imaging (biomechanical and morphological) data to estimate the risk of AAA rupture.[16,17] These models showed promising value in predicting clinical outcomes, including disease progression of AAA patients.

In the field of PAD, the majority of AI models have focused on disease detection and diagnosis. Predictive models for PAD disease progression and outcomes are less explored in this field.[18,19] Nonetheless, AI models for predicting in-hospital mortality[20] and long-term mortality[21,22] have been reported. Particularly, the model presented by Ross et al[21] considered diverse modalities including demographic, clinical, imaging, and genomic data to stratify PAD patients with a high risk of mortality. Similarly, the evaluation of major cardiovascular events presents a significant challenge in PAD. In this regard, models exist that predict MACE[21] or major adverse limb events (MALE).[23,24] A recent study by Li et al[24] has investigated MALE in a PAD patient population and achieved good model performance by combining demographic, clinical, and proteomic data in a multimodal AI model for risk prediction.

Despite the substantial progress in AI-powered diagnostics and risk stratification for AAA and PAD, key challenges remain. A significant hurdle is the lack of trustworthy AI development with large multimodal data sets and the actual integration of AI tools in clinical practice to aid clinical decision-making. Current models are frequently trained on data sets with a limited number of patients, potentially hindering their generalizability to more diverse populations and compromising external validity. Furthermore, the scope of existing models is often narrow and neglects the intricate interplay of factors contributing to AAA and PAD development, including genetic predispositions, lifestyle habits, and a broader spectrum of biomarkers. To achieve a more comprehensive understanding of these diseases, future AI models should prioritize multimodal data integration, incorporating information from diverse sources like electronic health records, imaging data, genetic data, and proteomic data. By addressing challenges and fostering robust, generalizable models in the vascular surgery field, AI has the potential to enhance the management of AAA and PAD.

The multidisciplinary European VASCUL-AID project is designed to overcome these challenges. The project aims to develop a platform for predicting disease progression and cardiovascular events in patients with AAA or PAD. The VASCUL-AID platform will support the creation of several AI models such as image analysis algorithms and risk prediction algorithms. In addition, the project includes the development of a mobile health app for patients and physicians to monitor their disease progression. In VASCUL-AID-RETRO, the retrospective study of this project, the aim is to develop predictive models based on data that has been previously collected. The study addresses key challenges in AI-powered diagnostics and risk stratification by utilizing a large multimodal patient data set. The inclusion of clinical and imaging data, alongside proteomics and genomics data, contributes to the comprehensive nature of the multimodal data set. Importantly, to achieve the
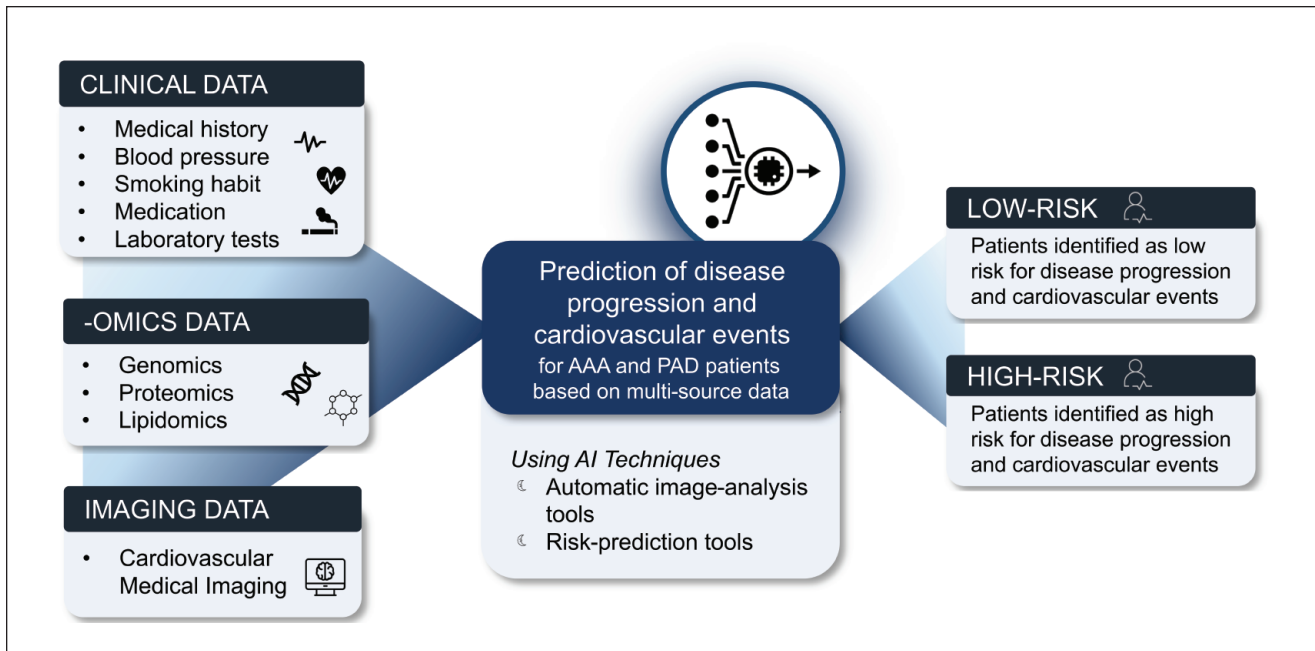
Figure 1. Overview of the VASCUL-AID-RETRO study.

establishment of trustworthy AI tools ready for the use in daily clinical care, ethical and legal considerations will be taken into account throughout the study. Through the extensive retrospective data set, the VASCUL-AID-RETRO study aims to develop trustworthy multimodal predictive AI models for multiple tasks including risk stratification of disease progression and MACE in patients with AAA and PAD. The developed models will be further adjusted by adding prospective data and will be internally validated in the future VASCUL-AID-PRO study.

Methods

Study Design and Study Population

The primary objective of this retrospective multicenter study is to develop and train algorithms to predict disease progression and risk of cardiovascular events in AAA and PAD patients by leveraging multimodal data from retrospective patient cohorts and biobanks. The secondary objective is to internally validate the developed algorithms using data from retrospective cohorts. An overview of the VASCUL-AID-RETRO study is shown in Figure 1.

For the development of the VASCUL-AID algorithms, patient data from electronic health records (EHRs) from European clinical consortium partners is available. The clinical partners of the VASCUL-AID consortium are Amsterdam University Medical Center (NL), Asklepios Kliniken Hamburg (DE), Centro Hospitalar e Universitário Sâo João (PT), Helsinki University Hospital (FI), Oxford University Hospitals (UK), and Serbian Clinical Center

(RS). All participating centers have obtained approval by ethical committees or other research bodies.

In this retrospective study, imaging and (reported) clinical EHR data of 5000 AAA and 6000 PAD patients will be collected from European consortium partners. Inclusion and exclusion criteria of patients are depicted in Table 1. Furthermore, blood samples of a subset of patients will be extracted from existing biobanks to perform multi-omics analyses. In addition, this consortium has access to large cohorts and registries with data for the envisioned study that will be used to enrich the retrospectively-collected VASCUL-AID data. All this data will be used to develop and train the algorithms for the predictions of cardiovascular events and AAA or PAD progression. An overview of the complete identified database including registries and biobanks that will be used in this study is described in Table 2. During the study, other data from new collaborations could be added as input to the prediction models.

During the entire VASCUL-AID-RETRO study, ethical and legal issues involved in developing trustworthy AI tools will be addressed. A Health Ethical Legal and Social Aspects (HELSA) framework of AI will be established early in the project and updated annually, to ensure compliance with evolving international laws and standards for responsible AI development. In this framework, relevant ethical, legal, and social issues and concerns for the VASCUL-AID-RETRO study will be discussed, and integrated, if necessary through guidelines. The HELSA issues include the fairness and inclusivity of AI (avoiding bias), data privacy, and exercising autonomy by patients. Within the HELSA context, potential liability risks of introducing

Table 1. Inclusion and Exclusion Criteria for VASCUL-AID-RETRO Abdominal Aortic Aneurysm (AAA) and Peripheral Arterial Disease (PAD) Study Cohorts.

AAA cohort

Inclusion criteria	Exclusion criteria

Male and female patients, aged 40-90 years, with:
- Any type of AAA with a diameter of 3.0 cm or larger, including:
    ○ Infrarenal, juxtarenal/pararenal, suprarenal/paravisceral/Crawford type IV thoracoabdominal aneurysms
    ○ AAAs on which previous intervention was performed
    ○ AAAs that ruptured
    ○ Inflammatory/infected aneurysms/penetrating aortic ulcers; *and*
- A visit to the vascular surgeon from 2015 to 2024
30% of the total database should be female patients.
75% of the patients must be preoperative AAA patients without inflammatory or infected aneurysms, or penetrating aortic ulcers, at baseline.
	Patients with only ascending, thoracic, thoracoabdominal (Crawford types I-III) aneurysms will be excluded.

PAD cohort

Inclusion criteria	Exclusion criteria

Male and female patients, aged 40-90 years, with:
- Lower extremity PAD, referring to atherosclerotic obstruction from the aorto-iliac segments to the pedal arteries, including:
    ○ Fontaine stages I-IV (including patients with critical limb ischemia)
    ○ Patients with acute limb ischemia
    ○ Patients who have had previous PAD intervention; *and*
- A visit to the vascular surgeon from 2015 to 2024
30% of the total database should be female patients.
75% of the patients must be claudicants, and not patients with critical limb-threatening ischemia or acute limb ischemia, at baseline.
	None.

Table 2. Identified Databases and Biobanks for VASCUL-AID-RETRO.

Database or biobank	Data specifications
Retrospectively-collected VASCUL-AID data	Clinical and imaging data of 5000 AAA and 6000 PAD patients.
Parel and LIVE Biobanks, Amsterdam University Medical Center	Clinical and imaging data of 250 AAA patients. Blood and tissue samples of 250 AAA and 50 PAD patients.
Kaiser Permanente, San Francisco Medical Center	Clinical and imaging data of 8000 AAA patients.
FinnGen Biobank, Finland	Estimation of clinical and -omics data of 4000 AAA and 18 000 PAD patients.
UK Biobank, the United Kingdom	Estimation of clinical and -omics data of 6000 AAA and 2000 PAD patients.
HELIUS database, Amsterdam University Medical Center	Estimation of clinical data of 100 AAA and 300 PAD patients from various population groups in the region of Amsterdam.

AI tools in PAD and AAA care will be investigated, followed by recommendations on how to avoid these risks and what safeguards are needed in terms of use and oversight.

Sample Size Calculation

To identify important predictive features for disease progression and cardiovascular risk, clinical input variables for AAA and PAD will be collected. The amount of clinical input variables will depend on a clinical variable list that will be established through expert consensus. These clinical variables will be enriched with imaging and multi-omics variables. To get to the final input parameters for the predictive AI models, principle component analyses or feature selection will be performed to select the maximum amount of variables that satisfy the rule of at least 10 events per predictor parameter[25] depending on the data distribution within the cohort of 5000 AAA and 6000 PAD patients.

Data Collection VASCUL-AID-RETRO Database

A clinical variable list will be established through expert consensus to select relevant variables from a clinical and ethical perspective and to minimize selection bias. Ethical considerations will be addressed through the involvement of ethical AI specialists that have assessed the variable list in accordance with the European AI Act that has been adopted in 2024. For the VASCUL-AID-RETRO database, clinical variables and imaging data will be collected from multiple follow-up visits with a vascular surgeon between 2015 until the latest reported available data in 2024. In addition, relevant EHR data from other specialties (such as cardiology, internal medicine, and nephrology) will be consulted to collect information about comorbidities and overall health status.

To collect the data, a data infrastructure is established that facilitates systematic data collection from all six participating clinical centers. The data infrastructure incorporates Castor EDC[26] for clinical data collection and the open-source eXtensible Neuroimaging Archive Toolkit (XNAT)[27] version 1.7.5 hosted by Health-RI[28] for imaging data storage and retrieval. The Castor EDC structure will specifically be designed for the VASCUL-AID-RETRO study and will incorporate the clinical variables determined with the expert consensus. The VASCUL-AID consortium partners will directly input their EHR data into Castor. Data from external databases and registries will be mapped to the VASCUL-AID clinical variables to enable integration of this data into Castor. Consequently, all data from different centers and external databases will be harmonized using the Castor designed for VASCUL-AID as a blueprint to obtain a uniform data set. Imaging of the VASCUL-AID consortium partners, as well as imaging from external databases, will be uploaded to the XNAT. The AI researchers will extract the data from these platforms for analyses.

For the multi-omics analyses, blood, plasma, and tissue samples of 250 AAA and 50 PAD patients will be collected from the existing Parel biobank and LIVE biobanks in Amsterdam University Medical Center. Blood plasma and tissue samples will be shipped to consortium partners in Porto for proteomics and lipidomics analyses. Targeted protein-resolved analyses will be performed to compare and validate the protein and lipid findings from blood and tissue samples, ensuring a comprehensive correlation between the two types of samples. To uncover the contribution of genetic variants related to AAA, PAD, and cardiovascular events, whole blood samples will be sent to consortium partners, VINS, in Serbia for genome-wide analysis. Genotyping will be performed using Infinium Global Screening Array-24+ v3.0, Illumina and scanned on Illumina iScan platform. This array provides extensive coverage (654 027 fixed genetic markers) and was commonly used in genome-wide association studies and meta-analyses.[29] Further implementation of imputation methods can enhance the resolution and accuracy of genotype data, enabling comprehensive genetic analyses.

Figure 2 illustrates the overview of the VASCUL-AID-RETRO data collection infrastructure.

Proposed Analyses

To obtain a comprehensive variable list, the consortium will collaboratively develop a VASCUL-AID key opinion leader (KOL) consensus-based clinical variable list. The initial list will be set up based on the literature[30,31] and shared with clinical and ethical VASCUL-AID consortium partners and external KOLs for evaluation. To reach a consensus on variables and definitions, consensus meetings will be held to discuss which variables and definitions should be included. It is expected that two expert consensus meetings will be sufficient to reach agreement. In case two physical meetings will not be sufficient, the remaining outcomes will be discussed through mailing. Extensive discussions involving physicians involved in the treatment of these patients, such as vascular surgeons, vascular internal medicine specialists, and cardiologists, will cumulate in a finalized list of variables. The KOLs will be from the VASCUL-AID consortium as well as external specialists in the field. The main categories of the consensus-based variable list will comprise demographics, vascular disease-related information, cardiac history, comorbidities, social/economic status, substance abuse, general fitness, vital parameters, family history, blood test and microbiology results, medication, diagnostic tests, and imaging data.

In addition to the variable list, an expert consensus-based core outcome set will be developed as part of the VASCUL-AID project. This outcome set will include clinical and patient-centered core outcome measures for AAA and PAD patients, covering outcomes related to general health or disease status, disease progression, such as components of MACE or MALE, treatment and treatment complications, or general disease status. To obtain a long list of potential outcomes, a systematic literature review and focus groups with patients and health care workers will be performed. Second, relevant stakeholders (patients, patient representatives, and health care workers) from consortium countries will be invited to participate in the Delphi study, which will conclude with a consensus meeting to determine the final core outcome sets, for AAA and PAD separately.

Following the collection of imaging, multi-omics data, and (reported) clinical EHR patient data of the AAA and PAD patients, baseline characteristics will be assessed using descriptive statistics. Results will be reported as proportions with percentages for categorical variables, and as mean with standard deviation or as median with interquartile range for continuous variables, as appropriate. Categorical data will be presented in numbers and percentages. Continuous variables will be presented with means
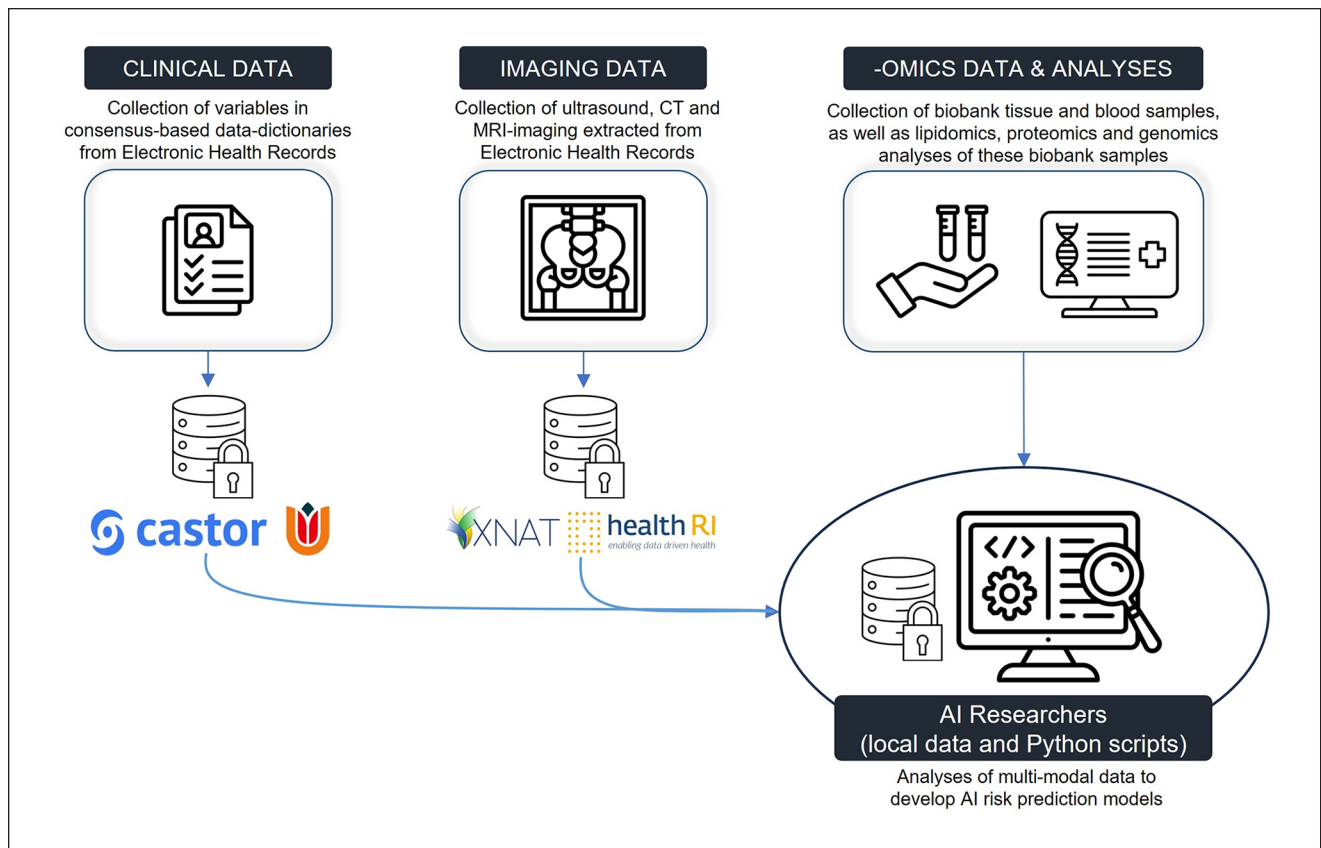
Figure 2. Schematic overview of VASCUL-AID-RETRO data infrastructure with Castor as the electronic data capturing system for health record data and an XNAT server from Health-RI for storage of imaging data.

and standard deviations or 95% confidence intervals or medians and the interquartile range, depending on their distribution. Survival analyses such as Kaplan-Meier estimation and Cox proportional hazards models will be performed to evaluate cardiovascular events and disease progression of the retrospective data set. The data set will comprise repeated measurements over time to perform longitudinal data analysis techniques, such as mixed-effects models and generalized estimating equations. Statistical analysis will be performed using R, SPSS, and Python.

Modality-specific analyses will be performed next. Automated image segmentation models will be developed using images collected from the following imaging modalities: computed tomography (CT), ultrasound (US), and magnetic resonance imaging (MRI). These models will be used to segment structures relevant to AAA and PAD, such as the vascular geometry, artery lumen, thrombus, calcifications, and perivascular fat. The AI models will then be trained on the segmented AAA and PAD artery geometries to estimate hemodynamic parameters such as pressure, velocity, and wall shear stress. Furthermore, mass spectroscopy-based proteomic and lipidomic analysis of plasma samples will be undertaken to identify differentially

secreted proteins and classify them into high vs low disease progression. Targeted genotyping will be done to identify the single nucleotide polymorphism information at previously identified risk loci. To help identify key variables or imaging features associated with cardiovascular events or disease progression, feature selection techniques may be utilized, such as recursive feature elimination or lasso regression (L1 regularization), in the development of the prediction models.

Initially, risk prediction models will be developed for each modality type (imaging, multi-omics, and reported clinical variables) separately. Various machine learning algorithms will be employed to create these individual prediction models. Finally, clinical data from the medical patient files, anatomical and hemodynamic data from medical imaging, and multi-omics data from patient blood samples will be integrated to be used as input to multimodal prediction models. Multiple AI models based on general machine learning (eg, random forest, XGBoost) or deep learning models (eg, convolutional neural networks, graph neural networks) will be assessed for optimal performance using suitable metrics (eg, sensitivity, area under the receiver operating curve) for its application. In the future

VASCUL-AID-PRO study, these developed AI models will be further refined and internally validated using prospectively collected data. All results will be published in open-source vascular or data-/AI-related journals, and the algorithms will be made publicly available on GitHub.

Risks and Challenges

Even though multimodal AI models have emerged as an effective method for studying and analyzing heterogeneous data, multiple challenges remain. Ethical and legal challenges arise with the development of the AI models including data security, privacy, handling data, and ethical regulations and policies per country.[32] These challenges should be overcome to ensure trustworthy analysis and interpretation. In this project, emphasis is placed on maintaining transparency concerning data sources, collection methods, and preprocessing steps in adherence to relevant laws, regulations, and ethical standards. Transparency and adherence to relevant regulations are ensured through collaboration with our consortium partners specialized in this field, who will set up the HELSA framework for the VASCUL-AID project. As data could be imbalanced, heterogeneous, or even missing at all,[32,33] it is important to review the VASCUL-AID data set and adjust it for imbalances that may arise. Also, data quality should be ensured by assessing the completeness and reliability of the data set.

Multiple techniques will be utilized to try to minimize the risks and to overcome these challenges. By utilizing a consensus-based variable list, most important variables will be included. Moreover, a single harmonized database is created in Castor EDC to store homogenous data sets from all participating partners. An Out-of-Distribution (OOD) filter will be implemented to identify and exclude patients who significantly deviate from the general population in our retrospective cohort. Implementing an OOD filter ensures that the data set used for the development of the models is representative of the target population, increasing the integrity and robustness of the models. Furthermore, to handle missing data techniques such as XGBoost and graph representation learning will be utilized.[34] Moreover, within the VASCUL-AID project, a study will be performed that will focus on the explainability of the developed algorithms through visualization to make the models as transparent as possible.

Conclusion

The VASCUL-AID-RETRO study will develop multimodal AI models, combining clinical, imaging, and multi-omics data, to predict AAA and PAD progression and associated risk of cardiovascular events in these patients for the first time. Following the VASCUL-AID-RETRO study, prospective VASCUL-AID studies will start in 2025 to collect prospective data and to test the algorithms in a pilot-setting (VASCUL-AID-PRO study). By integrating multimodal data sets, the final AI models have the potential to provide accurate predictions and more insights into disease progression, aiming to personalize treatments for improved patient outcomes.

ORCID iDs

Lotte Rijken https://orcid.org/0009-0008-9271-1868
Sabrina Zwetsloot https://orcid.org/0009-0009-8828-3108
Stefan Smorenburg https://orcid.org/0000-0002-4754-3016
Jelmer Wolterink https://orcid.org/0000-0001-5505-475X
Henk Marquering https://orcid.org/0000-0002-1414-6313
Roosmarie Jessen https://orcid.org/0009-0006-1353-5524
Katarzyna Bera https://orcid.org/0000-0003-3795-6762
Marina Dias-Neto https://orcid.org/0000-0002-7934-2016
Fabien Lareyre https://orcid.org/0000-0002-6765-8021
Igor Koncar https://orcid.org/0000-0002-4873-2775
Tamara Djuric https://orcid.org/0000-0002-9857-9828
Maarit Venermo https://orcid.org/0000-0001-8814-0988
Christian-Alexander Behrendt https://orcid.org/0000-0003-0406-3319
Noeska Smit https://orcid.org/0000-0002-3719-4625
Marlies Schijven https://orcid.org/0000-0001-7013-0116
Vincent Jongkind https://orcid.org/0000-0003-0129-1607
Venkat Ayyalasomayajula https://orcid.org/0000-0002-0158-6723
Kak Khee Yeung https://orcid.org/0000-0002-8455-286X

References

1. Golledge J. Abdominal aortic aneurysm: update on pathogenesis and medical treatments. *Nat Rev Cardiol*. 2019;16(4):225–242. doi:10.1038/s41569-018-0114-9.
2. Kullo IJ, Rooke TW. Clinical practice. Peripheral artery disease. *N Engl J Med*. 2016;374(9):861–871. doi:10.1056/NEJMcp1507631.
3. Kent KC. Clinical practice. Abdominal aortic aneurysms. *N Engl J Med*. 2014;371(22):2101–2108. doi:10.1056/NEJMcp1401430.
4. Chaikof EL, Brewster DC, Dalman RL, et al. The care of patients with an abdominal aortic aneurysm: the Society for

Vascular Surgery practice guidelines. *J Vasc Surg*. 2009;50:S2–S49. doi:10.1016/j.jvs.2009.07.002.

5. Darling RC, Messina CR, Brewster DC, et al. Autopsy study of unoperated abdominal aortic aneurysms. *Circulation*. 1977;56(3 suppl):II161–II164.

6. Bellamkonda KS, Nassiri N, Sadeghi MM, et al. Characteristics and outcomes of small abdominal aortic aneurysm rupture in the American College of Surgeons national surgical quality improvement program database. *J Vasc Surg*. 2021;74(3):729–737. doi:10.1016/j.jvs.2021.01.063.

7. Malyar N, Fürstenberg T, Wellmann J, et al. Recent trends in morbidity and in-hospital outcomes of in-patients with peripheral arterial disease: a nationwide population-based analysis. *Eur Heart J*. 2013;34(34):2706–2714. doi:10.1093/eurheartj/eht288.

8. Bath MF, Gokani VJ, Sidloff DA, et al. Systematic review of cardiovascular disease and cardiovascular death in patients with a small abdominal aortic aneurysm. *Br J Surg*. 2015;102(8):866–872. doi:10.1002/bjs.9837.

9. Swaminathan A, Vemulapalli S, Patel MR, et al. Lower extremity amputation in peripheral artery disease: improving patient outcomes. *Vasc Health Risk Manag*. 2014;10:417–424. doi:10.2147/vhrm.S50588.

10. Golledge J. Update on the pathophysiology and medical treatment of peripheral artery disease. *Nat Rev Cardiol*. 2022;19(7):456–474. doi:10.1038/s41569-021-00663-9.

11. Berger A, Simpson A, Leeper NJ, et al. Real-world predictors of major adverse cardiovascular events and major adverse limb events among patients with chronic coronary artery disease and/or peripheral arterial disease. *Adv Ther*. 2020;37:240–252. doi:10.1007/s12325-019-01132-z.

12. Tomee SM, Bulder RMA, Meijer CA, et al. Excess mortality for abdominal aortic aneurysms and the potential of strict implementation of cardiovascular risk management: a multifaceted study integrating meta-analysis, national registry, and PHAST and TEDY trial data. *Eur J Vasc Endovasc Surg*. 2023;65(3):348–357. doi:10.1016/j.ejvs.2022.11.019.

13. Saratzis A, Jaspers NEM, Gwilym B, et al. Observational study of the medical management of patients with peripheral artery disease. *Br J Surg*. 2019;106(9):1168–1177. doi:10.1002/bjs.11214.

14. Raffort J, Adam C, Carrier M, et al. Artificial intelligence in abdominal aortic aneurysm. *J Vasc Surg*. 2020;72(1):321–333.e1. doi:10.1016/j.jvs.2019.12.026.

15. Lee R, Jarchi D, Perera R, et al. Applied machine learning for the prediction of growth of abdominal aortic aneurysm in humans. *EJVES Short Rep*. 2018;39:24–28. doi:10.1016/j.ejvssr.2018.03.004.

16. Kleinstreuer C, Li Z. Analysis and computer program for rupture-risk prediction of abdominal aortic aneurysms. *Biomed Eng Online*. 2006;5:19. doi:10.1186/1475-925X-5-19.

17. Chung TK, Gueldner PH, Aloziem OU, et al. An artificial intelligence based abdominal aortic aneurysm prognosis classifier to predict patient outcomes. *Sci Rep*. 2024;14:3390. doi:10.1038/s41598-024-53459-5.

18. Lareyre F, Lê CD, Ballaith A, et al. Applications of artificial intelligence in non-cardiac vascular diseases: a bibliographic analysis. *Angiology*. 2022;73(7):606–614. doi:10.1177/00033197211062280.

19. Flores AM, Demsas F, Leeper NJ, et al. Leveraging machine learning and artificial intelligence to improve peripheral artery disease detection, treatment, and outcomes. *Circ Res*. 2021;128:1833–1850. doi:10.1161/CIRCRESAHA.121.318224.

20. Zhang D, Li Y, Kalbaugh CA, et al. Machine learning approach to predict in-hospital mortality in patients admitted for peripheral artery disease in the United States. *J Am Heart Assoc*. 2022;11:e026987. doi:10.1161/JAHA.122.026987.

21. Ross EG, Shah NH, Dalman RL, et al. The use of machine learning for the identification of peripheral artery disease and future mortality risk. *J Vasc Surg*. 2016;64(5):1515–1522. doi:10.1016/j.jvs.2016.04.026.

22. Callegari S, Romain G, Cleman J, et al. Long-term mortality predictors using a machine-learning approach in patients with chronic limb-threatening ischemia after peripheral vascular intervention. *J Am Heart Assoc*. 2024;13:e034477. doi:10.1161/JAHA.124.034477.

23. McBane RD, Murphree DH, Liedl D, et al. Artificial intelligence of arterial Doppler waveforms to predict major adverse outcomes among patients evaluated for peripheral artery disease. *J Am Heart Assoc*. 2024;13:e031880. doi:10.1161/JAHA.123.031880.

24. Li B, Nassereldine R, Zamzam A, et al. Development and evaluation of a prediction model for peripheral artery disease-related major adverse limb events using novel biomarker data. *J Vasc Surg*. 2024;80(2):490–497. doi:10.1016/j.jvs.2024.03.450.

25. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441. doi:10.1136/bmj.m441.

26. BV C. *Castor Electronic Data Capture*. Amsterdam: Castor; 2016.

27. Marcus DS, Olsen TR, Ramaratnam M, et al. The extensible neuroimaging archive toolkit: an informatics platform for managing, exploring, and sharing neuroimaging data. *Neuroinformatics*. 2007;5(1):11–34. doi:10.1385/ni:5:1:11.

28. Health RI. XNAT Service. Published 2024. Accessed January 16, 2025. https://www.health-ri.nl/en/services/xnat

29. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature*. 2021;600(7889):472–477. doi:10.1038/s41586-021-03767-x.

30. Oderich GS, Forbes TL, Chaer R, et al. Reporting standards for endovascular aortic repair of aneurysms involving the renal-mesenteric arteries. *J Vasc Surg*. 2021;73(1 suppl):4S–52S. doi:10.1016/j.jvs.2020.06.011.

31. Behrendt CA, Bertges D, Eldrup N, et al. International consortium of vascular registries consensus recommendations for peripheral revascularisation registry data collection. *Eur J Vasc Endovasc Surg*. 2018;56(2):217–237. doi:10.1016/j.ejvs.2018.04.006.

32. Char DS, Abràmoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. *Am J Bioeth*. 2020;20(11):7–17. doi:10.1080/15265161.2020.1819469.

33. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. *NPJ Digit Med*. 2022;5(1):171. doi:10.1038/s41746-022-00712-8.

34. You J, Ma X, Ding D, et al. *Handling Missing Data With Graph Representation Learning*. Vancouver, Canada: NeurIPS; 2020.