Decoding vascular age from brachial waveform morphology via machine learning and model-based data augmentation

IOP Publishing Mach. Learn.: Health 1 (2025) 015006 https://doi.org/10.1088/3049-477X/adf74f

MACHINE LEARNING
Health

OPEN ACCESS

RECEIVED
16 May 2025

REVISED
13 July 2025

ACCEPTED FOR PUBLICATION
4 August 2025

PUBLISHED
14 August 2025

PAPER

Decoding vascular age from brachial waveform morphology via machine learning and model-based data augmentation

Arian Aghilinejad*, Alessio Tamborini and Morteza Gharib

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA, United States of America
* Author to whom any correspondence should be addressed.

E-mail: aghili@caltech.edu

Keywords: blood pressure, aging, machine learning, Fourier decomposition, hemodynamics, biomedical signal processing

Supplementary material for this article is available online

Abstract

Vascular aging is increasingly recognized as a critical marker of cardiovascular health, yet no gold-standard method exists for its quantification. In this study, we present a machine learning (ML)–based framework for decoding vascular age (VasAge) from brachial pressure waveforms, under the assumption that vascular and chronological age are equivalent in healthy individuals. The ML model was trained on Fourier-based harmonic features extracted from high-resolution brachial waveforms obtained using a laboratory-developed cuff device. We tested our method on a clinical dataset of 111 subjects (45 women, mean age 65 years), consisting of 16 healthy and 95 unhealthy individuals. To develop the ML model, waveforms from the healthy subgroup were augmented using a synthetic waveform database generated by a physiologically relevant pulse wave computational model. For model development in the healthy population where chronological and vascular age are equivalent, the model's predictions correlated with measured age ($r = 0.91$), with marginal systemic bias (bounded to 5% of the mean). We then applied this model to the population of unhealthy individuals, finding higher vascular age than chronological age in this group (mean difference of 9.5 years), with sensitivity to elevated aortic stiffness and systolic blood pressure ($p < 0.05$). These results demonstrate that vascular age can be accurately estimated from the morphology of a single brachial pressure waveform using a Fourier-based ML approach, even with a moderately sized sample. This method offers a cost-effective, non-invasive strategy for personalized cardiovascular risk assessment and holds promise for clinical translation.

1. Introduction

Cardiovascular disease profoundly impacts individuals in the United States, across all stages of life and constitutes 17% of overall national health expenditures [1]. Age stands out as the primary determinant of cardiovascular health, with age significantly influencing risk factors for cardiovascular disease [2]. The ongoing demographic transition poses a substantial challenge for healthcare systems, with the proportion of the population aged over 65 set to double from 12% in 2010 to 22% by 2040 [3]. Although existing preventive strategies target individuals classified as high-risk based on conventional cardiovascular disease risk factors, these algorithms frequently overlook others who may also be at risk of developing cardiovascular disease. By identifying these individuals before clinical symptoms appear and implementing earlier detection and intervention, there is potential to decrease mortality and morbidity associated with cardiovascular disease [4, 5]. Consequently, there exists a pressing imperative to innovate new methods and devices for the diagnosis and monitoring of cardiovascular diseases. Preferably, these methods should rely on noninvasive measurements to mitigate medical complications and enhance patient comfort, while also being user-friendly for both medical professionals and individuals in home settings [6–8].

Developing informative markers for cardiovascular disease requires a thorough understanding of the underlying mechanisms within the cardiovascular system. The circulatory system functions through a finely

tuned hemodynamic equilibrium involving the heart, vascular network, and major physiological system [9, 10]. This optimal hemodynamic coupling can be disrupted by cardiovascular disease [11, 12]. From the physical point of view, the rhythmic beating of the heart generates blood pressure fluctuations that propagate as waves through the arterial network, with reflections occurring at impedance mismatch points, typically characterized by junctions in arterial structure and changes in vessel elasticity [13, 14]. This pulsatile blood flow dictates various aspects of circulatory physiology and pathology, with pulsatility manifesting in pressure waves across the entire circulatory system [13, 15, 16]. The comprehension of arterial wave propagation mechanisms has long been a focal point in hemodynamics, with waves carrying considerable diagnostic potential that remains underutilized [5]. Traditional clinical practice primarily concentrates on fundamental pulse pressure features, such as systolic and diastolic pressure, for diagnosis, prevention, and treatment. However, recent years have seen an increasing recognition of other wave characteristics and finer pulse details achievable through analyzing the shape of the pressure waveform [5, 17–22]. These considerations have led to the emergence of various analytical techniques, such as pulse wave analysis [20, 23–25], wave intensity analysis [17, 26–28], and wave separation analysis [18, 29–31]. With advancements in noninvasive pressure measurement techniques, a rise in the utilization of wave analysis methods and wave-based information in future clinical practices is foreseeable.

Motivated by the significance of chronological age in the development of cardiovascular disease, there has been a recent effort to introduce a new concept known as vascular age by quantifying the age of such vessels [32–35]. When used relative to an individual's chronological age these can provide valuable insight about the aging process of the cardiovascular system. However, the direct quantification of vascular age remains underdeveloped [36]. As an example, Mitchell *et al* [19] employed a deep learning-based approach to first derive carotid-femoral pulse wave velocity (a measure of aortic stiffness) from a single pressure waveform and then used linear regression to rescale the derived pulse wave velocity to estimate vascular age. However, this approach is based on the assumption that there is a linear relationship between age and wave speed. In this study, we approach the problem of defining and quantifying vascular age from an engineering and data science perspective. Our core assumption is that, in a healthy population, vascular age (VasAge) corresponds to chronological age, and that the information embedded within the pressure waveform can be decoded to reveal this VasAge. This way, we propose an alternative method that does not restrict model development to pulse wave velocity or any other specific physiological parameter or cardiac marker. To achieve this, we incorporate a machine learning approach designed to extract physics-based information encoded in the pressure signal [30, 37, 38], train the model on a healthy population, and then use it to decode vascular age in a samples of healthy and unhealthy subjects. This approach represents a shift in perspective regarding vascular age, creating a direct link with the morphology of the pressure waveform.

The central hypothesis of this study is that VasAge is encoded in the morphology of the arterial pressure waveform, and that an appropriate combination of signal decoding and ML techniques can extract and quantify this information. Our primary objective is to develop a method that decodes VasAge from brachial pressure waveforms using wave-based signal analysis integrated with ML, and to evaluate the clinical relevance of the decoded VasAge, particularly its association with cardiovascular conditions such as arterial stiffening. ML is a strong candidate for this task due to its ability to model complex, nonlinear relationships within physiological signals, enabling the extraction of subtle age-related features that may not be accessible through traditional analytical methods. To test this hypothesis, we utilized data from an in-house investigational study employing a laboratory-developed cuff device capable of high-resolution pulse waveform acquisition. To address the limitations posed by the modest sample size of the clinical dataset, we systematically augmented the training set using a virtual population generated from a well-established computational model of arterial pulse wave propagation, a common practice in machine learning, particularly when clinical data are difficult to acquire at scale. The resulting vascular age model and the associated decoding strategy provide a promising framework for personalized cardiovascular assessment.

2. Materials and methods

2.1. Clinical dataset

We analyzed brachial pressure waveforms collected during an investigational study involving a high-resolution cuff-based acquisition device, as previously described [39, 40]. The study cohort included both healthy individuals, recruited for proof-of-concept testing, and patients who had undergone surgical intervention. The cuff was positioned on the participant's left arm in accordance with standard placement guidelines. Exclusion criteria were applied at two levels: (i) patient-level exclusions included severe post-operative cardiac events following left heart catheterization, inability to obtain a brachial blood pressure measurement, and contraindications to cardiac catheterization as determined by the interventional cardiologist. (ii) Signal-level exclusions were based on quality control procedures and included cuff or

Table 1. Baseline characteristics of patient data ($N = 111$).

Variable	Healthy ($N = 16$)	Unhealthy ($N = 95$)
Age, years	67 ± 11	65 ± 8
Women, N (%)	10 (63%)	35 (37%)
Height, cm	167 ± 9	171 ± 10
Weight, kg	69 ± 9	88 ± 18
Heart rate, bpm	70 ± 14	69 ± 10
Blood pressure, mmHg		
Systolic	119 ± 11	128 ± 16
Diastolic	72 ± 10	77 ± 12

catheter malfunctions, device saturation beyond calibration limits, measurement errors, severe arrhythmias, poor signal quality, or algorithmic failures. After applying these criteria, the final dataset comprised 111 individuals. To estimate aortic stiffness, we calculated pulse arrival time as the time difference, in milliseconds, between the peak of the ECG R-wave and the foot of the brachial pressure waveform acquired by the cuff device. This value was then normalized by subject height to derive a surrogate pulse wave velocity (sPWV). While this method is an approximation, since pulse arrival time includes the pre-ejection period and height is only a rough proxy for arterial path length, it provides a practical estimate of vascular stiffness in our dataset. For this study, which focuses on associating pressure waveform features with vascular age, this level of approximation is considered sufficient.

The study protocol was approved by the Institutional Review Boards of Western and Salus, and all participants provided written informed consent prior to enrollment. The study conformed to the ethical principles outlined in the Declaration of Helsinki. The cuff device used in this study was an investigational tool for high-resolution, non-invasive pulse waveform acquisition. It incorporated a blood pressure module (NIBP 2020 UP) with oscillometric and tourniquet modes, as well as a custom pneumatic system for capturing pulse waveforms. Data were acquired at a hold pressure corresponding to the subject's mean arterial pressure and recorded continuously for 20 s. This pressure was selected to align more closely with the simulated waveforms described in the following section. Waveform data were digitized at a sampling rate of 1 kHz. To develop an age-encoded model of healthy vascular function, we isolated a subset of healthy participants based on the absence of the following conditions: hypertension (systolic BP $\geqslant$ 140 mmHg or diastolic BP $\geqslant$ 90 mmHg), diabetes, hyperlipidemia, cardiovascular disease, current smoking, or obesity (BMI $\geqslant$ 30 kg m^{-2}) [41]. The characteristics of the healthy versus non-healthy participants are summarized in table 1.

2.2. Healthy virtual dataset

To address the limitations of the moderately sized clinical dataset for machine learning development, we employed a computationally generated virtual dataset for training purposes (i.e. data augmentation). Since our training was restricted to healthy individuals, where chronological age and vascular age were assumed to be equivalent, we utilized a virtual cohort representing healthy aging [42]. This dataset was derived from a validated one-dimensional (1D) arterial pulse wave propagation model [43]. In this model, the arterial tree is represented as a network of 1D segments, with blood flow and pressure dynamics governed by conservation of mass and momentum applied to each control volume [44, 45]. The governing equations, expressed in terms of cross-sectional area A and flow rate Q, are given by

$$\begin{pmatrix} \frac{\partial A}{\partial t} \\ \frac{\partial Q}{\partial t} \end{pmatrix} + \begin{pmatrix} \frac{\partial Q}{\partial x} \\ \frac{\partial}{\partial x}\left(\alpha \frac{Q^2}{A}\right) + \frac{A}{\rho}\frac{\partial P}{\partial x} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{f}{\rho} \end{pmatrix}. \tag{1}$$

Here, α is shape factor accounting for the velocity profile, P is the intraluminal pressure, and f represents the frictional term [43]. To close this system, a Voigt-type viscoelastic wall model was adopted [42, 43]. The aortic inflow was prescribed as a periodic waveform, and three-element Windkessel model was implemented at the outlet boundary of peripheral arteries. Using this physiologically informed framework and parameter values confined to healthy ranges, an open-access virtual dataset was generated previously by Charlton *et al* [42]. For the purposes of this study, we selected a subset of virtual subjects aged 50 years and older who had normal blood pressure (systolic $<$ 130 mmHg, diastolic $<$ 80 mmHg), normal cardiac output ($>$4 l min^{-1}), and physiological carotid-femoral pulse wave velocity ($<$14 m s^{-1}). This yielded a final cohort of 794 virtual subjects, which was used to augment the healthy training dataset. The characteristic of the virtual dataset is presented in table S1 of the supplementary material. To evaluate the physiological relevance of the virtual dataset used for augmenting the healthy clinical data, we compared key pressure waveform features between

the virtual and clinical healthy subsets (table S2). As shown in the table, the mean values of several waveform features are closely aligned across the two groups. We also observed that the distribution of feature values is wider in the clinical dataset, which is expected due to real-world variability. To provide visual context, figures S1 and S2 in the supplementary material present representative waveform examples from both the virtual and clinical healthy populations.

2.3. Data handling and preprocessing

For the purpose of model development in this study, we partitioned the data to isolate the healthy population, under the assumption that vascular age is equivalent to chronological age in this group. The objective was to develop a model that relates the non-invasive pressure waveform to vascular age, which, in the case of healthy individuals, corresponds directly to their actual age. We augmented the healthy clinical dataset with a computationally generated virtual dataset. Each pressure waveform in the healthy cohort was associated with the individual's chronological age and used as a single training sample. This yielded a total of 1148 waveforms, consisting of 354 signals from 16 healthy clinical participants and 794 signals from the virtual aging dataset. This combined dataset was used for training and validating the healthy-age encoded regression model. The model's performance was then evaluated by applying it to the remaining 95 clinical subjects with condition to estimate their vascular age. For the healthy cohort, waveform segmentation was performed using R-peak to R-peak intervals from the ECG, and although no explicit correction was applied for breathing-induced fluctuations, the model implicitly accounted for these variations by learning from multiple beats; for the unhealthy group, a representative time-averaged waveform was used. This single representative waveform was derived by averaging individual beat-to-beat segments. Each beat was first resampled to a common length, determined by the median number of samples across all beats, and then a pointwise average was computed to obtain the final waveform. To calibrate the cuff-derived waveforms, we applied a previously validated method that corrects raw systolic blood pressure (SBP) and diastolic blood pressure values based on oscillometric readings [40]. The corrected systolic and diastolic blood pressure values were used to define the peaks and troughs of individual cardiac cycles, and fluctuations were incorporated using the dynamic envelope calibration method [40, 46].

To transform each waveform from its high-dimensional time-series form into a compact feature space for machine learning, we evaluated two fundamental signal decomposition strategies: (i) Fourier Decomposition into harmonics, and (ii) empirical mode decomposition (EMD). Fourier decomposition represents a periodic signal as a sum of sinusoids at different frequencies, termed harmonics. It assumes the signal is stationary over the interval of interest. The Fourier series decomposition is commonly defined as:

$$P(t) = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi}{T} nt\right) + b_n \sin\left(\frac{2\pi}{T} nt\right) \right), \tag{2}$$

where T is the period of the pressure function $P(t)$ (i.e. the cardiac cycle or inverse of heart rate for blood pressure waveform). In conducting the Fourier decomposition here, we used the sinusoid-cosinusoide form to extract the coefficients at different frequencies for constructing input features for regression-based machine learning [37, 39]. Coefficients a_n and b_n are associated with each individual harmonics (cosine and sine) corresponding to different frequencies $f_n = \frac{n}{T}$, and can be calculated by the Fourier transform given by:

$$a_n = \frac{2}{T} \int_0^{\mathrm{T}} P(t) \cos\left(\frac{2\pi}{T} nt\right) \mathrm{d}t, \tag{3}$$

$$b_n = \frac{2}{T} \int_0^{\mathrm{T}} P(t) \sin\left(\frac{2\pi}{T} nt\right) \mathrm{d}t, n \in (0, \infty). \tag{4}$$

By obtaining the coefficients using equations (3) and (4), the represented pressure waveform $\tilde{P}(t)$ with a finite number of selected frequencies (i.e. $n = 0$ to N) is constructed by summing each individual frequency component (equation (2)). In this study, the pressure waveform features were extracted as the first five to twenty low-frequency components using the fast Fourier transform to capture the overall contour shape of the waveform [37, 38, 47]. In contrast to Fourier analysis, EMD is a data-adaptive and non-parametric technique designed for analyzing physiological signals [48, 49]. EMD decomposes a signal into a finite set of oscillatory components known as intrinsic mode functions (IMFs). In this regard, the signal $P(t)$ can be expressed as

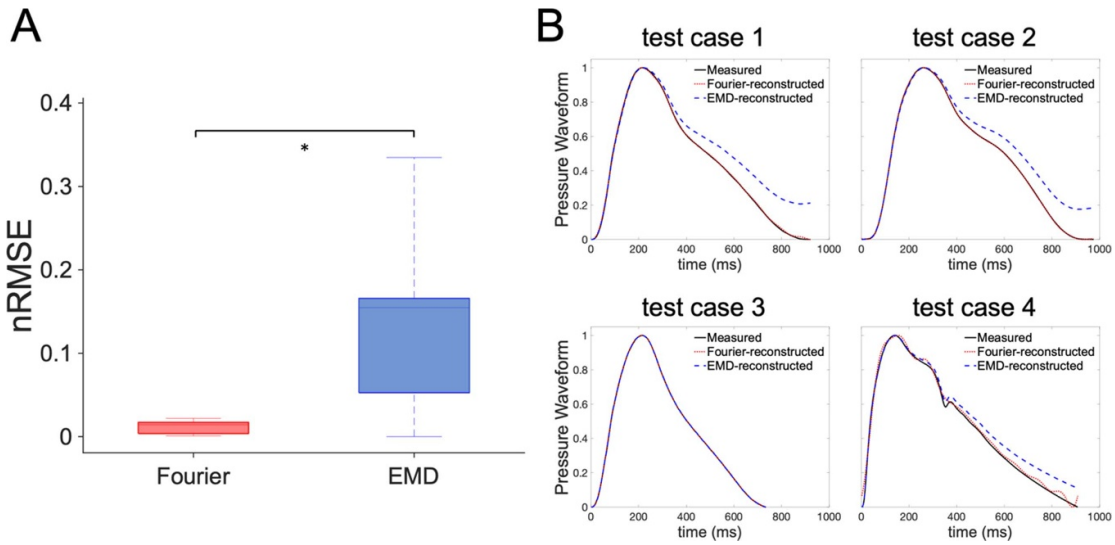

Figure 1. (a) Statistical analysis of the associated error between the original (measured) and reconstructed pressure waveform from Fourier-based decomposition and empirical mode decomposition (EMD). (b) Four sample cases of the measured and reconstructed pressure waveform from Fourier-based decomposition and empirical mode decomposition (EMD). For Fourier-reconstruction, 10 harmonics is used.

$$P(t) = \sum_{i=1}^{K} \mathrm{IMF}_i(t) + r(t), \tag{5}$$

where $\mathrm{IMF}_i(t)$ denotes the ith IMF, and $r(t)$ is the residual trend. The decomposition is obtained through an iterative process called sifting, which isolates IMFs from the highest to lowest frequency content. In our framework, we applied EMD to each waveform and extracted the first few IMFs to capture localized frequency and amplitude features associated with vascular dynamics. Figure 1 compares the reconstruction accuracy of the two methods. Figure 1(a) shows the distribution of normalized root-mean-square error (nRMSE) between the original and reconstructed waveforms. Figure 1(b) provides four representative examples of original waveforms overlaid with reconstructions using both Fourier and EMD approaches. As illustrated, the Fourier-based method consistently outperformed EMD in capturing the overall shape and oscillatory behavior of pressure waveforms. While EMD offers the advantage of adaptively capturing local features, it exhibited greater reconstruction error and variability across subjects. Due to its higher fidelity in waveform reconstruction and compatibility with regression-based models, we selected Fourier decomposition as the primary method for feature extraction in this study.

2.4. Machine learning model

Machine learning model development was carried out using 80% of the healthy cohort, combining both clinical and virtual datasets. This subset was used exclusively for model training, while the remaining 20% of healthy subjects was reserved as a validation set for evaluating the model's ability. In this study, vascular age was defined as equivalent to chronological age for the healthy population, providing ground truth labels for supervised learning. Each pressure waveform was preprocessed using Fourier decomposition, and a fixed number of harmonics were extracted as features. Specifically, both sine and cosine coefficients from the first N harmonic components were used to construct the input feature vector. These features capture the frequency-domain characteristics of the waveform and serve as a compact representation of its morphology. The input of the machine learning model consists of a vector of sine and cosine numeric coefficients derive from Fourier decomposition of the brachial pressure along with the heart rate, and the output of the machine learning model is the chronological age in the healthy subset.

The model architecture consisted of a fully connected feedforward neural network with an expansion-contraction design. It included five hidden layers with 32, 64, 128, 64, and 32 neurons, respectively. All layers employed the ReLU activation function, and the output layer was linear to predict age as a continuous variable. The model was trained using the Adam optimizer with a learning rate of 0.001 for 1000 epochs. A mean squared error loss function was used to minimize the difference between predicted and true age values during training. To assess the model's robustness, we systematically varied the number of Fourier harmonics used as input features and evaluated the sensitivity of the model to different frequency resolutions. All training and testing procedures were performed in Python using the TensorFlow framework

for model implementation. Data preprocessing, including waveform normalization (scaling the raw signal between zero and one, and then calibrate) and Fourier transformation, was conducted using NumPy and Pandas in Python (Python Software Foundation, Python Language Reference, version 3.11). To avoid any potential confusion regarding data leakage, we emphasize that the virtual dataset was split prior to model development such that each virtual individual was used exclusively in either training or testing. Additionally, the key evaluation was performed solely on the unhealthy clinical dataset, which was never seen during training and serves as an independent hold-out test set.

2.5. Vascular age decoding framework

An overview of the proposed framework for decoding vascular age (VasAge) is presented in figure 2(a). The central concept of this framework is the construction of a healthy-age encoded model, which serves as a personalized decoder for estimating vascular age from a single non-invasive pressure waveform. This model is trained exclusively on a healthy population, under the assumption that their vascular and chronological ages are equivalent. The resulting system enables inference of vascular age by mapping waveform features to age labels learned from the healthy cohort. The core components of the analytical pipeline used to construct this model are outlined in figure 2(a). Raw brachial pressure waveforms, recorded continuously over a 20 s interval, are segmented on a per-subject basis. Each waveform is calibrated using the procedure described in the previous section. The virtual dataset was split into non-overlapping training and testing subsets prior to model development. Each waveform in the virtual dataset was generated using a unique combination of physiological parameters, ensuring that no identical waveform appeared in both the training and testing sets. Model assessment on the unhealthy clinical population was performed once using a completely held-out test set, ensuring a rigorous and unbiased evaluation. To reduce the dimensionality of the pressure signal and capture its dominant frequency components, we apply Fourier-based signal decomposition. Figure 1 demonstrates the method's ability to preserve key morphological features with reduced input dimensionality. The resulting low-dimensional representation of each pressure waveform serves as input to a trained deep neural network, described previously. The output of this model is the estimated age of the individual, assuming they are healthy, based solely on their pressure waveform. This predicted value constitutes the VasAge. To quantify vascular risk, we define the VasAge Risk as the difference between the model-estimated vascular age and the subject's actual chronological age (VasAge-Age). A positive VasAge Risk indicates that the vasculature appears 'older' than expected for the individual's age, while a negative value suggests 'younger' vascular function. Once trained, the healthy-age encoded model is applied to the unhealthy clinical cohort to estimate their VasAge, using only a single brachial pressure waveform as input (figure 2(b)). Figure S3 in the supplementary material presents the training loss history as a function of epoch during the training for the healthy-age encoded model.

2.6. Statistical analysis

To assess the accuracy of the healthy-age encoded model, we evaluated Pearson correlation coefficient (r). The agreement and bias between the computed and measured variables were assessed using Bland-Altman analysis, which presents mean differences with limits of agreement (mean bias $\pm$ 1.96 standard deviations of the differences). The Kruskal–Wallis rank-sum test was used to assess the vascular age risk as a function of SBP and arterial stiffness (quantified by surrogate pulse wave speed). Statistical significance was defined as $p < 0.05$. All mathematical and statistical analyses of the clinical data were performed using custom-written codes implemented in Python.

3. Results

3.1. Data processing workflow

Figure 3(a) illustrates the data processing workflow used to construct the healthy subset for developing the healthy-age encoded model, as well as the independent datasets for determining VasAge in both healthy and unhealthy subjects. A total of 1,243 brachial pressure waveforms were analyzed, including 449 clinical waveforms acquired via cuff measurements and 794 computationally generated waveforms. The initial clinical cohort of 111 individuals was divided into 16 healthy participants (providing 354 waveforms) and 95 high-risk individuals who did not meet one or more of the health-associated criteria. An 80:20 train-test split was applied to create independent cohorts: one for model development and the other as a validation test group for VasAge evaluation. Note that model development was conducted exclusively on the healthy sample. Figure 3(b) shows the chronological age distribution of both the healthy (clinical and virtual) and unhealthy groups used in this study. The similarity in age distributions across groups supports a fair evaluation of the healthy-age encoded model for VasAge decoding. Additionally, figure S4 in the supplementary material

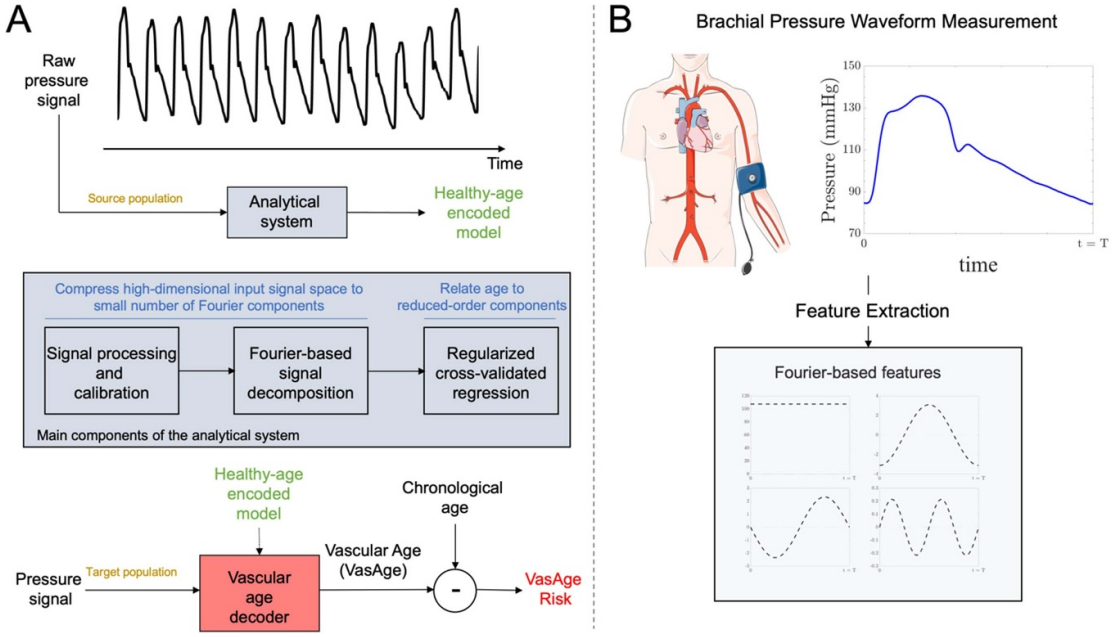

Figure 2. (a) An overview of the decoding process for vascular age. The first step in constructing this decoder is to develop an analytical system that generates a healthy-age encoded model. Main components to develop the healthy-age encoded model is also demonstrated. (b) The input feature extraction to construct healthy-age encoded model from single generic brachial pressure waveform measurement based on Fourier-based and time-based components.

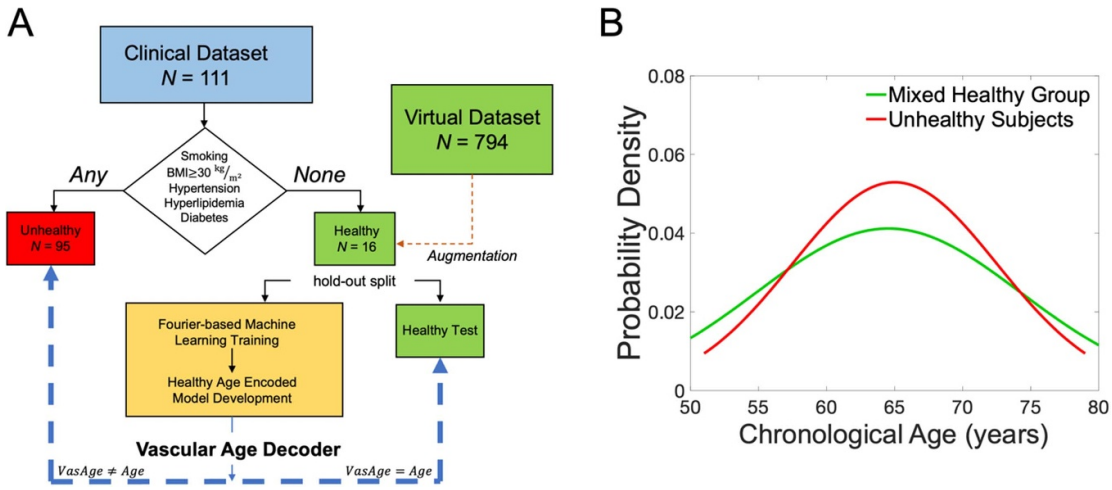

Figure 3. (a) Data processing workflow used to create the healthy subset for developing the healthy-age encoded model, as well as the independent datasets for vascular age (VasAge) determination, which include both healthy and unhealthy samples. (b) Probability density of chronological age for the mixed healthy group (including both clinical and virtual subjects) and the unhealthy clinical group.

presents the age distributions of each subgroup within the healthy dataset, including both clinical and virtual subjects for reference.

3.2. Vascular age determination

To determine VasAge across different populations, we first evaluate the accuracy of the healthy-age encoded model. This model is trained and validated exclusively on a healthy subgroup, using brachial pressure waveforms to estimate age. Table 2 summarizes the agreement between the model-predicted age and the actual chronological age, including both error metrics and Bland–Altman analysis. The model employed is a fully connected neural network that uses input features derived from the Fourier decomposition of the signal into its fundamental harmonics. The data in table 2 are from a validation set within the healthy subgroup, which was used only once for model assessment. The table also examines how the number of input Fourier harmonics affects prediction accuracy. Results suggest that using 10 input harmonics provides a good

Table 2. Statistics on the testing healthy sample as a function of the number of input Fourier harmonics.

Model	Correlation coefficient	RMSE	Mean diff (LoA)
5 input Fourier Harmonics	0.90	5.77	4.03 (−4.07, 12.13)
10 input Fourier Harmonics	0.91	4.87	2.93 (−4.70, 10.57)
15 input Fourier Harmonics	0.91	4.63	2.00 (−6.19, 10.20)
20 input Fourier Harmonics	0.85	5.97	2.27 (−8.55, 13.09)

RMSE stands for root mean square error and has the units of years. LoA stands for limit of agreement. Mean Diff and LoA has units of years.

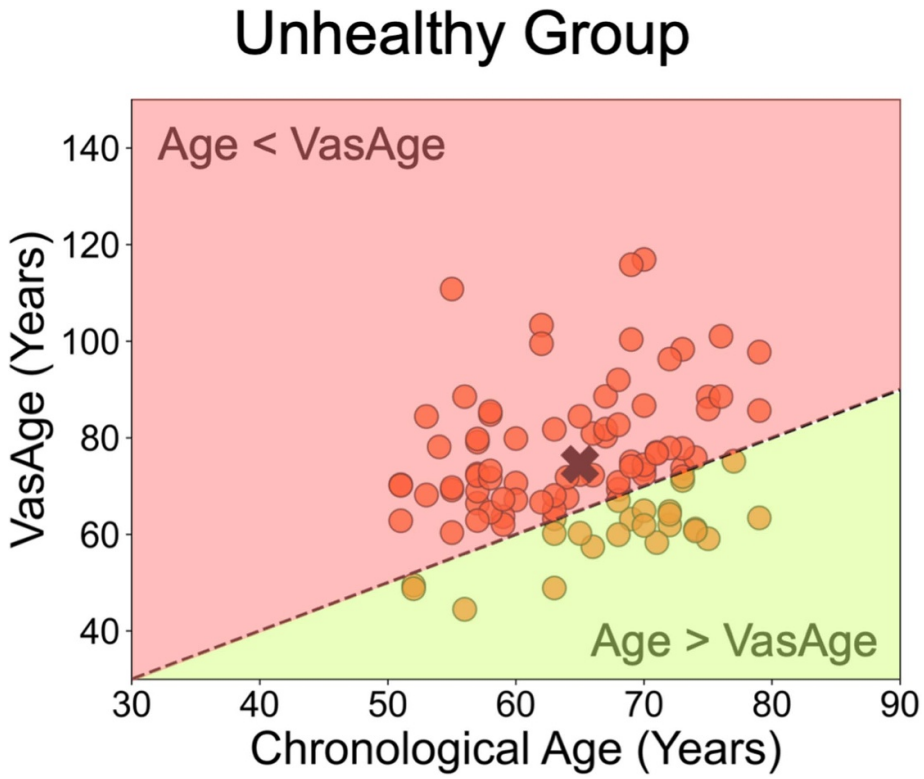

Figure 4. Scatter plots comparing vascular age (VasAge) with chronological age in the held-out unhealthy group. The VasAge is derived from the Fourier-decomposed features extracted from the brachial pressure waveform.

balance between model accuracy and input dimensionality. Notably, the correlation coefficient between the predicted age, obtained through machine learning from a single brachial waveform, and chronological age is 0.91. According to the Bland-Altman analysis, the mean difference between predicted and actual age is 2.93 years (less than 5% of the average age in the healthy group), with limits of agreement ranging from −4.70 to 10.57 years.

Figure 4 presents scatter plots comparing VasAge and chronological age in the held-out unhealthy individuals. VasAge values are generated using the healthy-age encoded model applied to brachial pressure waveforms from this group. For reference, the results of applying this model to the validation sample consisting of healthy subjects are shown in figure S5 of the supplementary material. In the healthy group, most data points lie near the identity line (VasAge $\sim$ Age), indicating strong agreement. In contrast, the unhealthy group shows a marked shift, with many data points falling above the identity line (VasAge $>$ Age), suggesting vascular aging beyond chronological age. Cross mark indicates the mean difference of VasAge and chronological age for the sample group. The mean difference between VasAge and chronological age is 2.9 years in the healthy group and 9.5 years in the unhealthy group, corresponding to 4% and 15% of their respective mean ages.

3.3. Physiological-relevancy of the vascular age

Figure 5 illustrates the comparison of VasAge risk, defined as the difference between VasAge and chronological age, across subgroups stratified by sPWV and SBP. In figure 5(a), subjects are grouped based on sPWV, which is calculated from height-adjusted pulse arrival time. Figure 5(b) presents a similar analysis using SBP categories. This analysis is performed exclusively on the unhealthy clinical dataset (hold-out test

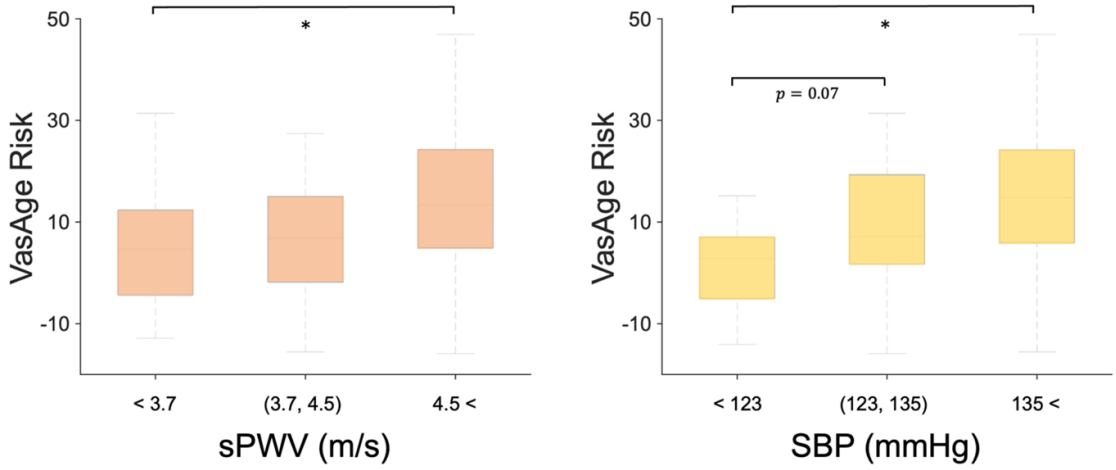

Figure 5. Boxplots for vascular age (VasAge) risk distributions for different levels of aortic stiffness quantified by surrogate pulse wave velocity (sPWV) and systolic blood pressure (SBP). $^{*}p < 0.05$.

set) to explore how vascular aging patterns vary with established cardiovascular risk markers. The results show that individuals with elevated aortic stiffness, as indicated by higher sPWV values, tend to have both higher VasAge and higher chronological age compared to those with lower aortic stiffness (figure S6 and S7 of the supplementary material). More importantly, the VasAge risk is significantly greater in the high-stiffness group ($p < 0.05$), indicating that vascular aging may accelerate disproportionately relative to chronological aging as aortic stiffness increases. This suggests that VasAge captures pathophysiological changes in vascular structure and function that may not be fully explained by age alone. In parallel, when subjects are grouped by SBP (figure 5(b)), a similar trend is observed. Individuals with higher SBP exhibit elevated VasAge and chronological age compared to those with lower SBP (figure S8 of the supplementary material). Additionally, VasAge risk is significantly greater in the high-SBP group ($p < 0.05$), suggesting that increased systolic pressure is also associated with an accelerated vascular aging process.

4. Discussion

In this study, we developed a ML–based framework to estimate VasAge from a single brachial pressure waveform. The method was evaluated on a clinical dataset that included both healthy and unhealthy individuals. A key finding is that VasAge, derived solely from waveform morphology, serves as a sensitive marker of cardiovascular health and shows a strong association with risk factors such as arterial stiffness. Our approach extracts reduced-order features from the waveform using harmonic decomposition, which are then used to train the ML model. To address the limited size of the clinical dataset, we augmented the training data with virtual samples generated from a physiologically validated pulse wave model. These results demonstrate that the proposed signal-processing and ML pipeline enables accurate vascular age estimation, even in data-constrained clinical settings. This highlights the potential of ML to extract physiologically meaningful insights from high-dimensional biomedical signals that are otherwise difficult to interpret through conventional methods.

Recently, there has been a significant shift towards analyzing the elements within the shape of the pressure waveform [5, 13, 19, 20]. Building on these studies, we aimed to leverage the entire brachial pressure waveform as input to our health assessment algorithm for determining VasAge. We developed a healthy-age encoded model trained exclusively on a healthy cohort, where VasAge is assumed to match chronological age. In this model, Fourier-based features extracted from pressure waveforms, acquired using a laboratory-developed cuff device, served as inputs to a machine learning algorithm. The model was trained to estimate each individual's chronological age, which corresponds to VasAge in the healthy population. To evaluate model performance, we compared predicted age with actual chronological age in an independent validation set that was kept blind throughout model development and used only once for evaluation (table 2). Statistical analysis revealed strong agreement between predicted and measured age, indicating high accuracy of the model. The use of decomposition enabled transformation of complex waveform contours into constituent harmonic components that reflect forward and reflected waves propagating through the arterial system. These Fourier-based features provided a physiologically meaningful and compact input space for the neural network model. Because the model was trained solely on a healthy cohort, its output

corresponds to VasAge, representing the age-equivalent health status inferred from waveform morphology. The results in table 2 further highlight that the number of Fourier input harmonics affects predictive accuracy. Specifically, using 10 harmonics offered a favorable balance between performance and input dimensionality, minimizing overfitting. While increasing the number of input Fourier harmonics initially enhances model performance, we observe diminishing returns and even a decline in accuracy beyond 15 harmonics. This trend likely reflects the fact that higher-order harmonics begin to capture noise and measurement artifacts rather than physiologically relevant waveform features. Including more harmonics introduces high-frequency components that contribute little to physiological interpretation but increase the risk of overfitting, ultimately compromising model robustness.

Figure 4 shows results from applying the healthy-age encoded model to new, previously unseen sample of unhealthy population. Applying the healthy-age encoded model to this new group decoded the brachial pressure waveform to reveal its VasAge. For the healthy population, we expect a one-to-one relationship between VasAge and chronological age, which is evident in figure S5 of the supplementary material. Conversely, for the high-risk population, we expect a larger VasAge compared to chronological age. As shown in figure 4, individuals in the unhealthy group are predominantly located above the line representing a one-to-one relationship (the region where VasAge $>$ Age). The elevated VasAge in the unhealthy population demonstrates the model's ability to differentiate between health and disease conditions based on a single brachial pressure waveform.

We also examined the associations of VasAge with aortic stiffness by categorizing participants into groups with sPWV < 3.7 m s^{-1}, 3.7–4.5 m s^{-1}, and >4.5 m s^{-1} (figure 5). Results indicate that VasAge risk (the difference between VasAge and chronological age) shows significant differences among these groups. Previous research at Framingham and elsewhere has demonstrated that pulse wave velocity is strongly related to the progression of various risk factors and an increased risk of several adverse clinical outcomes [4, 15, 50, 51]. These findings motivated researchers to define VasAge mainly based on carotid-femoral pulse wave velocity. In a recent population-based study, Mitchell *et al* [19] proposed an approach to estimate VasAge based on carotid-femoral pulse wave velocity derived from applying convolutional neural networks to a single pressure waveform. In this study, we proposed an alternative approach that does not constrain our model development to pulse wave velocity or any other physiological parameter or cardiac markers. To address how we define VasAge, we partition data and assume that vascular and chronological ages are equivalent in a healthy, low-risk population. Our underlying hypothesis is that the information contained in the pressure waveform can be decoded to reveal VasAge. By using wave decomposition of the pressure waveform, we first decode VasAge in the healthy population (healthy-age encoded model) and then apply it to other samples. Results from figure 5 show that elevated VasAge is associated with increased aortic stiffness, which reflects the natural connections within the cardiovascular system. A similar trend is observed for SBP, where VasAge risk is significantly associated with higher SBP levels.

Our study has several strengths and potential limitations. One limitation is the imbalance in population size between the healthy and unhealthy subsets. While our findings suggest that the proposed approach is sensitive to an individual's health status, future studies should incorporate larger and more balanced datasets. Further investigations could also examine the associations between VasAge and clinical outcomes, including coronary heart disease, cardiovascular events, heart failure, and extracardiac conditions such as cerebrovascular incidents. The age distribution in our study population reflects the older individuals with a higher risk of cardiovascular disease. While this aligns with our goal of developing a diagnostic marker for at-risk populations, future studies should evaluate the applicability of the proposed method in more diverse cohorts. A core assumption of this study is that, within the healthy subgroup, vascular age aligns with chronological age. While this assumption provides practical value for model development, it may introduce bias, as vascular aging is influenced by a complex interplay of genetic, environmental, and lifestyle factors not captured in our dataset. Even in the absence of overt cardiovascular disease, individuals may exhibit different rates of vascular aging depending on factors such as physical activity, diet, stress levels, sleep patterns, and genetic background. Future studies that include these variables could help refine this assumption and enhance the physiological accuracy of vascular age estimation models. In addition, while carotid-femoral pulse wave velocity is widely used to estimate vascular stiffness and correlate it with age, such data were not available in our dataset; future work using large population-based datasets could enable direct comparison with established vascular aging methods. Despite these limitations, our study has notable strengths. We employed a reduced-order representation of the input pressure waveform, which improves model generalizability and is suitable for smaller datasets. Given the demonstrated ability of our reduced-order model to accurately reconstruct the full waveform, we do not anticipate significant performance gains from using deep learning with the full signal. Moreover, our approach offers a more data-efficient solution, which is especially advantageous when working with moderate-sized clinical datasets. Our method-derived vascular

age effectively distinguishes between healthy and unhealthy subjects and is sensitive to elevated arterial stiffness and SBP.

5. Conclusion

In summary, we have introduced a method for predicting a clinically significant VasAge that is closely linked to the occurrence of cardiovascular events such as arterial stiffening. This approach has been validated in a clinical sample with brachial pulse derived from the laboratory-developed cuff device. The VasAge model offers a novel strategy for personalized medicine, facilitating both the assessment and long-term monitoring of cardiovascular health. This method is noninvasive, relying solely on the analysis of a single brachial pressure waveform, which makes it suitable for clinical implementation with relatively low-cost equipment and minimal training requirements. While current preventive measures focus on individuals identified as high-risk through traditional cardiovascular disease risk factors, these algorithms often miss individuals who are at risk of developing cardiovascular disease. By identifying these individuals before clinical symptoms emerge and providing earlier detection and intervention, it may be possible to reduce cardiovascular disease-related mortality and morbidity.

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

Data were obtained from a data transfer and use agreement between Caltech and Avicena LLC (d.b.a. Ventric Health). Figure 2 was partly generated with a modified figure from Servier Medical Art, provided by Servier, licensed under a Creative Commons Attribution 3.0 unported license.

Conflict of interest

The authors report no conflict of interest.

Author contributions

Arian Aghilinejad 0000-0001-5231-427X
Conceptualization (lead), Data curation (equal), Formal analysis (lead), Investigation (lead), Methodology (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (equal)

Alessio Tamborini
Conceptualization (supporting), Data curation (equal), Formal analysis (supporting), Investigation (supporting), Methodology (supporting), Writing – review & editing (equal)

Morteza Gharib
Conceptualization (supporting), Investigation (supporting), Resources (lead), Supervision (lead), Writing – review & editing (equal)

References

[1] Arnett D K, Blumenthal R S, Albert M A, Buroker A B, Goldberger Z D, Hahn E J, Himmelfarb C D, Khera A, Lloyd-Jones D and McEvoy J W 2019 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: executive summary: a report of the American college of cardiology/American heart association task force on clinical practice guidelines *Circulation* **140** e563–95
[2] Oxenham H and Sharpe N 2003 Cardiovascular aging and heart failure *Eur. J. Heart Fail.* **5** 427–34
[3] North B J and Sinclair D A 2012 The intersection between aging and cardiovascular disease *Circ. Res.* **110** 1097–108
[4] Chirinos J A, Segers P, Hughes T and Townsend R 2019 Large-artery stiffness in health and disease: JACC state-of-the-art review *J. Am. Coll. Cardiol.* **74** 1237–63
[5] Cooper L L, Rong J, Pahlevan N M, Rinderknecht D G, Benjamin E J, Hamburg N M, Vasan R S, Larson M G, Gharib M and Mitchell G F 2021 Intrinsic frequencies of carotid pressure waveforms predict heart failure events: the Framingham heart study *Hypertension* **77** 338–46
[6] Charlton P H, Allen J, Bailón R, Baker S, Behar J A, Chen F, Clifford G D, Clifton D A, Davies H J and Ding C 2023 The 2023 wearable photoplethysmography roadmap *Physiol. Meas.* **44** 111001
[7] Pahlevan N M, Rinderknecht D G, Tavallali P, Razavi M, Tran T T, Fong M W, Kloner R A, Csete M and Gharib M 2017 Noninvasive iphone measurement of left ventricular ejection fraction using intrinsic frequency methodology *Crit. Care Med.* **45** 1115–20

[8] Chandrasekhar A, Kim C-S, Naji M, Natarajan K, Hahn J-O and Mukkamala R 2018 Smartphone-based blood pressure monitoring via the oscillometric finger-pressing method *Sci. Trans. Med.* **10** eaap8674

[9] Mitchell G F, van Buchem M A, Sigurdsson S, Gotal J D, Jonsdottir M K, Kjartansson Ó, Garcia M, Aspelund T, Harris T B and Gudnason V 2011 Arterial stiffness, pressure and flow pulsatility and brain structure and function: the age, gene/environment susceptibility–Reykjavik study *Brain* **134** 3398–407

[10] Aghilinejad A, Bilgi C, Geng H and Pahlevan N M 2025 Aortic stretch and recoil create wave-pumping effect: the second heart in the systemic circulation *J. R. Soc. Interface* **22** 20240887

[11] Maillard P, Mitchell G F, Himali J J, Beiser A, Tsao C W, Pase M P, Satizabal C L, Vasan R S, Seshadri S and DeCarli C 2016 Effects of arterial stiffness on brain integrity in young adults from the Framingham heart study *Stroke* **47** 1030–6

[12] Aghilinejad A, Wei H, Magee G A and Pahlevan N M 2022 Model-based fluid-structure interaction approach for evaluation of thoracic endovascular aortic repair endograft length in type B aortic dissection *Front. Bioeng. Biotechnol.* **10** 825015

[13] Van de Vosse F N and Stergiopulos N 2011 Pulse wave propagation in the arterial tree *Annu. Rev. Fluid Mech.* **43** 467–99

[14] Aghilinejad A, Amlani F, King K S and Pahlevan N M 2020 Dynamic effects of aortic arch stiffening on pulsatile energy transmission to cerebral vasculature as a determinant of brain-heart coupling *Sci. Rep.* **10** 1–12

[15] Mitchell G F, Hwang S-J, Vasan R S, Larson M G, Pencina M J, Hamburg N M, Vita J A, Levy D and Benjamin E J 2010 Arterial stiffness and cardiovascular events: the Framingham heart study *Circulation* **121** 505–11

[16] Aghilinejad A, Amlani F and Gharib M 2025 Power-frequency relationship of wave dynamics in fluid-filled compliant tubes *Phys. Rev. Fluids* **10** 033102

[17] Chiesa S T, Masi S, Shipley M J, Ellins E A, Fraser A G, Hughes A D, Patel R S, Khir A W, Halcox J P and Singh-Manoux A 2019 Carotid artery wave intensity in mid-to late-life predicts cognitive decline: the Whitehall II study *Eur. Heart J.* **40** 2300–9

[18] Cooper L L, Rong J, Benjamin E J, Larson M G, Levy D, Vita J A, Hamburg N M, Vasan R S and Mitchell G F 2015 Components of hemodynamic load and cardiovascular events: the Framingham heart study *Circulation* **131** 354–61

[19] Mitchell G F, Rong J, Larson M G, Korzinski T J, Xanthakis V, Sigurdsson S, Gudnason V, Launer L J, Aspelund T and Hamburg N M 2024 Vascular age assessed from an uncalibrated, noninvasive pressure waveform by using a deep learning approach: the AI-VascularAge model *Hypertension* **81** 193–201

[20] Mynard J P, Kondiboyina A, Kowalski R, Cheung M M and Smolich J J 2020 Measurement, analysis and interpretation of pressure/flow waves in blood vessels *Front. Physiol.* **11** 1085

[21] Tamborini A, Aghilinejad A and Gharib M 2024 Reconstructing invasive aortic pressure waveforms from non-invasive brachial measurements using a machine learning approach *Circulation* **150** A4141635

[22] Tamborini A and Gharib M 2024 Listening to heart sounds through the pressure waveform *Sci. Rep.* **14** 26824

[23] Pahlevan N M, Tavallali P, Rinderknecht D G, Petrasek D, Matthews R V, Hou T Y and Gharib M 2014 Intrinsic frequency for a systems approach to haemodynamic waveform analysis with clinical applications *J. R. Soc. Interface* **11** 20140617

[24] Aghilinejad A, Alavi R, Rogers B, Amlani F and Pahlevan N M 2021 Effects of vessel wall mechanics on non-invasive evaluation of cardiovascular intrinsic frequencies *J. Biomech.* **129** 110852

[25] Yofoglu L K, Karachalias F, Georgakis M K, Tountas C, Argyris A A, Zhang Y, Papaioannou T G, Blacher J, Weber T and Vlachopoulos C 2023 Association of pressure wave reflections with left ventricular mass: a systematic review and meta-analysis *Hypertension* **80** e29–e42

[26] Aghilinejad A, Amlani F, Liu J and Pahlevan N M 2021 Accuracy and applicability of non-invasive evaluation of aortic wave intensity using only pressure waveforms in humans *Physiol. Meas.* **42** 105003

[27] Parker K H 2009 An introduction to wave intensity analysis *Med. Biol. Eng. Comput.* **47** 175

[28] Kang J, Aghilinejad A and Pahlevan N M 2019 On the accuracy of displacement-based wave intensity analysis: effect of vessel wall viscoelasticity and nonlinearity *PLoS One* **14** e0224390

[29] Aghilinejad A and Gharib M 2024 Use of machine learning for wave separation analysis from non-invasive pressure measurement and its applicability in capturing aortic stiffening *J. Am. Coll. Cardiol.* **83** 4599

[30] Aghilinejad A and Gharib M 2024 Assessing pressure wave components for aortic stiffness monitoring through spectral regression learning *Eur. Heart J. Open* **4** oeae040

[31] Zamani P, Jacobs Jr D R, Segers P, Duprez D A, Brumback L, Kronmal R A, Lilly S M, Townsend R R, Budoff M and Lima J A 2014 Reflection magnitude as a predictor of mortality: the multi-ethnic study of atherosclerosis *Hypertension* **64** 958–64

[32] Qian Z, Huang Y, Zhang Y, Yang N, Fang Z, Zhang C and Zhang L 2024 Metabolic clues to aging: exploring the role of circulating metabolites in frailty, sarcopenia and vascular aging related traits and diseases *Front. Genet.* **15** 1353908

[33] Zhang C, Tao J and Cardiovascular Group S O G 2018 Chinese medical association, 2018, "expert consensus on clinical assessment and intervention of vascular aging in China (2018) *Aging Med.* **1** 228–37

[34] Consortium A B, Zhang L, Guo J, Liu Y, Sun S, Liu B, Yang Q, Tao J, Tian X-L and Pu J 2023 A framework of biomarkers for vascular aging: a consensus statement by the aging biomarker consortium *Life Med.* **2** lnad033

[35] Zanelli S, Agnoletti D, Alastruey J, Allen J, Bianchini E, Bikia V, Boutouyrie P, Bruno R M, Climie R and Djamaleddine D 2024 Developing technologies to assess vascular ageing: a roadmap from VascAgeNet *Physiol. Meas.* **45** 121001

[36] Gómez-Sánchez L, Tamayo-Morales O, Suárez-Moreno N, Bermejo-Martín J F, Domínguez-Martín A, Martín-Oterino J A, Martín-González J I, González-Calle D, García-García Á and Lugones-Sánchez C 2023 Relationship between the structure, function and endothelial damage, and vascular ageing and the biopsychological situation in adults diagnosed with persistent COVID (BioICOPER study). A research protocol of a cross-sectional study *Front. Physiol.* **14** 1236430

[37] Aghilinejad A, Tamborini A and Gharib M 2024 A new methodology for determining the central pressure waveform from peripheral measurement using Fourier-based machine learning *Artif. Intell. Med.* **154** 102918

[38] Aghilinejad A, Wei H and Pahlevan N M 2023 Non-invasive pressure-only aortic wave intensity evaluation using hybrid Fourier decomposition-machine learning approach *IEEE Trans. Biomed. Eng.* **70** 2139–48

[39] Tamborini A, Aghilinejad A and Gharib M 2025 A spectral machine learning approach to derive central aortic pressure waveforms from a brachial cuff *Proc. Natl Acad. Sci.* **122** e2416006122

[40] Tamborini A and Gharib M 2024 Validation of a suprasystolic cuff system for static and dynamic representation of the central pressure waveform *J. Am. Heart Assoc.* **13** e033290

[41] Mitchell G F, Parise H, Benjamin E J, Larson M G, Keyes M J, Vita J A, Vasan R S and Levy D 2004 Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham heart study *Hypertension* **43** 1239–45

[42] Charlton P H, Mariscal Harana J, Vennin S, Li Y, Chowienczyk P and Alastruey J 2019 Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes *Am. J. Physiol. Heart. Circ. Physiol.* **317** H1062–H85

[43] Alastruey J, Parker K H and Sherwin S J Arterial pulse wave haemodynamics *Proc. 11th Int. Conf. on Pressure Surges, Virtual PiE Led T/a BHR Group* pp 401–43
[44] Mynard J P and Smolich J J 2015 One-dimensional haemodynamic modeling and wave dynamics in the entire adult circulation *Ann. Biomed. Eng.* **43** 1443–60
[45] Aghilinejad A, Amlani F, Mazandarani S P, King K S and Pahlevan N M 2023 Mechanistic insights on age-related changes in heart-aorta-brain hemodynamic coupling using a pulse wave model of the entire circulatory system *Am. J. Physiol. Heart. Circ. Physiol.* **325** H1193–209
[46] Tamborini A and Gharib M 2023 A pneumatic low-pass filter for high-fidelity cuff-based pulse waveform acquisition *Ann. Biomed. Eng.* **51** 2617–28
[47] Aghilinejad A, Rogers B, Geng H and Pahlevan N M 2023 On the longitudinal wave pumping in fluid-filled compliant tubes *Phys. Fluids* **35** 091903
[48] Li C, Wu Y, Lin H, Li J, Zhang F and Yang Y 2022 ECG denoising method based on an improved VMD algorithm *IEEE Sens. J.* **22** 22725–33
[49] Roy B, Roy A, Chandra J K and Gupta R 2020 I-PRExT: photoplethysmography derived respiration signal extraction and respiratory rate tracking using neural networks *IEEE Trans. Instrum. Meas.* **70** 1–9
[50] King K S, Sheng M, Liu P, Maroules C D, Rubin C D, Peshock R M, McColl R W and Lu H 2018 Detrimental effect of systemic vascular risk factors on brain hemodynamic function assessed with MRI *Neuroradiol. J.* **31** 253–61
[51] Maroules C D, Khera A, Ayers C, Goel A, Peshock R M, Abbara S and King K S 2014 Cardiovascular outcome associations among cardiovascular magnetic resonance measures of arterial stiffness: the Dallas heart study *J. Cardiovasc. Magn. Reson.* **16** 1–9