

# SCIENTIFIC REPORTS



OPEN

## Artificial Intelligence Estimation of Carotid-Femoral Pulse Wave Velocity using Carotid Waveform

Peyman Tavallali<sup>1</sup>, Marianne Razavi<sup>1</sup> & Niema M. Pahlevan<sup>2,3,4</sup>

In this article, we offer an artificial intelligence method to estimate the carotid-femoral Pulse Wave Velocity (*PWV*) non-invasively from one uncalibrated carotid waveform measured by tonometry and few routine clinical variables. Since the signal processing inputs to this machine learning algorithm are sensor agnostic, the presented method can accompany any medical instrument that provides a calibrated or uncalibrated carotid pressure waveform. Our results show that, for an unseen hold back test set population in the age range of 20 to 69, our model can estimate *PWV* with a Root-Mean-Square Error (RMSE) of 1.12 *m/sec* compared to the reference method. The results convey the fact that this model is a reliable surrogate of *PWV*. Our study also showed that estimated *PWV* was significantly associated with an increased risk of CVDs.

Cardiovascular diseases (CVDs) and stroke are among the major causes of death in the United States and the total cost related to them was more than \$316 billion in 2011–2012<sup>1,2</sup>. New cardiovascular monitoring methods are urgently needed in order to limit the growing burden of CVDs. Arterial stiffening is one of the risk factors for CVDs<sup>3,4</sup>, which can be assessed non-invasively by calculating the carotid to femoral *PWV*<sup>5</sup>. This parameter is a gold standard of arterial stiffness, the rate at which pressure waves move down the aortic vessel<sup>6</sup>. Increased arterial stiffness is related to an increased risk of cardiovascular events; therefore, it has become an independent marker for CVDs<sup>6,7</sup>. Because of its clinical significance, there has been a surge in addressing arterial stiffness and *PWV*<sup>8</sup>. Arterial stiffness and its surrogates such as *PWV* have been suggested as one of the risk factors along with other biomarkers such as high cholesterol, diabetes, and left ventricular hypertrophy when cardiovascular risk is being evaluated<sup>8</sup>. Past studies have shown a strong correlation between *PWV* and the presence of CVDs<sup>9–14</sup>.

Although carotid-femoral *PWV* measurement is non-invasive, this process is intrusive as it requires the waveform collection from inguinal region. Obtaining accurate carotid-femoral *PWV* measurements often requires a well-trained staff within a clinical setting<sup>15</sup>. The need of the medical community is an easy-to-use and non-intrusive method to measure carotid-femoral *PWV* with acceptable accuracy and precision; see ref.<sup>16</sup>.

At the same time, recent advances in the field of artificial intelligence have opened up new areas and methods in creating novel modeling and predictive methods for clinical use<sup>17</sup>. The model and analysis in this paper are in accord to this path of introducing artificial intelligence to the field of medical sciences.

In this study, a novel, easy-to-use, and non-invasive approach to estimate carotid-femoral *PWV*, from a single carotid waveform measurement, is explored. This method is based on the newly developed Intrinsic Frequency (IF) algorithm<sup>18,19</sup>. IF method solely needs one uncalibrated trace of a carotid, or aortic, pressure waveform. Our method takes an uncalibrated trace of carotid pressure waveform and performs IF analysis and basic signal processing on it. Then, it combines the results of these analyses with easy-to-obtain clinical parameters such as age. Finally, it models the *PWV* by neural networks through bootstrap averaging. The main advantages in having an estimated *PWV* from an uncalibrated carotid pressure waveform, with few typical clinical variables such as blood pressure, would be that it does not need an ECG measurement nor a femoral tonometry recording. As a result, it is easier, and potentially can be done by a smart phone as we have shown in our previous publication<sup>20</sup>.

<sup>1</sup>Avicena, LLC, 2400 N Lincoln Ave, Altadena, CA, 91001, USA. <sup>2</sup>Department of Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA, USA. <sup>3</sup>Division of Cardiovascular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. <sup>4</sup>Huntington Medical Research Institutes, Advanced Imaging Center, Pasadena, CA, USA. Peyman Tavallali and Marianne Razavi contributed equally to this work. Correspondence and requests for materials should be addressed to P.T. (email: [tavallali@gmail.com](mailto:tavallali@gmail.com))

## Data Description

We used the Framingham Heart Study (FHS) data, a longitudinal epidemiological cohort analysis, in this manuscript<sup>21</sup>. The participants were part of FHS Cohorts Gen 3 Exam 1<sup>22</sup>, Offspring Exam 7<sup>23</sup>, and Original Exam 26<sup>24</sup>. They underwent a comprehensive, noninvasive assessment of central hemodynamics generating a successful collection of a total of  $N = 6698$  tonometry recordings. The recorded  $PWV$  measurements were calculated using a simultaneous right carotid tonometry pressure waveform with electrocardiogram recording and a right femoral tonometry pressure waveform with electrocardiogram recording combined with body surface measurements of the participants<sup>7</sup>. This method of measuring is sometimes called the sequential measurement<sup>25</sup>. One can find a comparison of this method to the reference method of measuring  $PWV$  in<sup>25</sup>. For a broader description regarding FHS data, please see<sup>7</sup> and references contained within.

Some participants had missing or erroneous tonometry waveforms data. We marked these recordings as “faulty record” ( $N = 1011$ ). From the rest of the data ( $N = 5687$ ), some recorded  $PWV$  measurements had values equal to or greater than  $30\text{ m/sec}$ , or even  $0\text{ m/sec}$ . We considered these values to be “measurement error” ( $N = 21$ ). We also excluded the population of age greater than 70 ( $N = 661$ ) to minimize the CVD treatment effects. Also, we know that individuals having an age of 70 or greater already experience arterial stiffness due to age factor. These filterings led to a total of  $N = 5020$  usable observations. Furthermore, for the prognostics study, we excluded individuals having cardiovascular diseases prior to, or on, their tonometry exam date leading to  $N = 4798$  participants.

## Artificial Intelligence Methodology

**General Signal Processing.** Each uncalibrated tonometry recording from the FHS data included a 10 to 20 second trace. Some of the signal processing parameters were provided by FHS data. For example the unitless variables Augmentation Index ( $AIx$ ) and Mean Carotid Shape Factor ( $MCSF$ ) were included as they were provided (See<sup>26</sup> and references contained within). As a quick reference,  $MCSF$  is the average value of an arterial cardiac cycle signal normalized by its range. Furthermore, Reflected Wave Arrival Time ( $RWAT$ ) was also among the variables that FHS had provided with the data (See<sup>26</sup> and references within).

However, to use the IF algorithm, we had to extract arterial cycles from the raw signal.

At first, a short-window moving average was used to eliminate the unwanted noise from the signals. The window size was taken to be  $0.02\text{ sec}$ . This window size would eliminate noise levels above  $\sim 50\text{ Hz}$ . The blood pressure waveform can be encoded with frequencies less than  $\sim 25\text{ Hz}$  ( $50\text{ Hz}$  is twice this value)<sup>27</sup>. In mathematical terms, the moving average  $\bar{s}(t)$  we used in our study, for a signal  $s(t)$ , can be expressed as

$$\bar{s}(t) = \frac{1}{0.04} \int_{t-0.02}^{t+0.02} s(\tau) d\tau. \quad (1)$$

Our method is not dependent on this choice of noise-filtering and other low-pass filtering approaches could also be used. Then the signals were normalized to remove the effects of breathing and other artificial motions. The normalization was performed using the location of maxima and minima of each recording<sup>28</sup>. We then used a modified version of the automatic cycle selection introduced in<sup>29</sup> to pick cycles. Dicrotic notch was found based on the derivatives and filtering of the picked cycles<sup>28</sup>. These cycles were then fed into the IF algorithm.

**Intrinsic Frequency.** A typical arterial pressure waveform consists of a systolic and diastolic part. The systolic part is when the aortic valve is open and heart is pumping blood into the aorta and arterial system. The diastolic part is when the aortic valve is closed preventing the blood from re-entering the left ventricle. The closure of the aortic valve, on the pressure waveform trace, is commonly called the dicrotic notch. The IF method<sup>18,19</sup> assumes that there are two constant dominant dynamical frequencies before and after the closure of the aortic valve. These frequencies are called Intrinsic Frequencies (IFs). This method does not need a calibrated aortic or carotid pressure signal; and can even be applied to signals collected by a smart phone<sup>20</sup>.

In the IF method<sup>18,19</sup>, it is assumed that the instantaneous frequencies are piecewise constant throughout the cardiac cycle. The dicrotic notch separates these frequencies. For an aortic pressure waveform, the IF problem can be formulated as

$$S(a_1, b_1, \bar{p}, \omega_1; t) = (a_1 \cos \omega_1 t + b_1 \sin \omega_1 t + \bar{p}) \mathbf{1}_{[0, T_0)}(t) + (a_2 \cos \omega_2 t + b_2 \sin \omega_2 t + \bar{p}) \mathbf{1}_{[T_0, T)}(t), \quad (2)$$

with a continuity condition at  $T_0$  (the time of the dicrotic notch) and periodicity at  $T$  (the duration of the cardiac cycle). Here, the *indicator function* is defined as

$$\mathbf{1}_{[x, y)}(t) = \begin{cases} 1, & x \leq t < y, \\ 0, & \text{else.} \end{cases} \quad (3)$$

Also,  $a_1$ ,  $b_1$ ,  $a_2$  and  $b_2$  are the envelopes of the IF model fit.  $\omega_1$  and  $\omega_2$  are Intrinsic Frequencies (IFs) of the waveform. Further,  $\bar{p}$  is the mean pressure for the period  $[0, T)$ . The goal of the IF model (2) is to extract a fit, called Intrinsic Mode Function (IMF), that carries most of the energy (information) from a pressure waveform  $s(t)$  in one period. The latter is done by solving the following optimization problem<sup>19</sup>:

$$\underset{a_1, b_1, \omega_1, \bar{p}}{\text{minimize}} \quad \|s(t) - S(a_1, b_1, \bar{p}, \omega_1; t)\|_2^2 \quad (4)$$

$$\text{subject to } \begin{aligned} a_1 \cos \omega_1 T_0 + b_1 \sin \omega_1 T_0 &= a_2 \cos \omega_2 T_0 + b_2 \sin \omega_2 T_0, \\ a_1 &= a_2 \cos \omega_2 T + b_2 \sin \omega_2 T. \end{aligned} \quad (5)$$

Here,  $\|\cdot\|_2$  is the  $L^2$ -norm defined on  $[0, T)$ . One assumption in this optimization is that the extracted IMF is continuous at the diastolic notch time  $T_0$  (the first condition in (5)). The other assumption is that the extracted IMF is periodic (the second condition in (5)). The method of the solution of the optimization problem mentioned by (4) and (5) can be found in<sup>19</sup>.

**Statistical Learning.** Dimensionless parameters or combinatorial mixtures can be extracted from the solutions of (4) and (5). These parameters are both of mathematical and physiological importance as shown in our recent work of noninvasive iPhone measurement of left ventricular ejection<sup>20</sup>. For example, we can normalize  $\omega_1$  and  $\omega_2$  with respect to systolic and diastolic periods,  $T_0$  and  $T - T_0$  respectively<sup>20</sup>. Even, we can normalize  $\omega_1$  and  $\omega_2$  by the whole cardiac cycle  $T$ . In fact, there is a systematic way to create new variables from a set of given features. We have used the method introduced in<sup>30</sup> to create new features from  $\omega_1, \omega_2, T$ , and  $T - T_0$ . Some of the outputs of this method were used in our *PWV* model.

The original set of variables used for feature extraction included IFs and their variants, carotid waveform shape factors such as reflected wave arrival time and augmentation index, and clinical features and blood pressure and age. We specifically used the mentioned clinical variables since a 2010 study published by European heart journal has emphasized that *PWV* is affected by age and blood pressure<sup>31</sup>. In short, the original set of features is

$$V_0 = \{\omega_1, \omega_2, \bar{\omega}_1, \bar{\omega}_2, \omega_{1c}, \omega_{2c}, \omega_{1n}, \omega_{2n}, \rho, E_r, C_r, T, \frac{1}{T - T_0}, \frac{1}{T_0}, \text{MCSF}, \text{AIx}, \text{SSN}, \text{RWAT}, \text{Age}, P_s, P_d\}. \quad (6)$$

In (6), we have constructed new features from  $\{\omega_1, \omega_2, T, T - T_0\}$ . To be more specific, we used the methodology introduced in<sup>30</sup> and field expert knowledge to introduce

$$\bar{\omega}_1 = \omega_1 T_0, \quad (7)$$

$$\bar{\omega}_2 = \omega_2 (T - T_0), \quad (8)$$

$$\omega_{1c} = \omega_1 \sqrt{T_0}, \quad (9)$$

$$\omega_{2c} = \omega_2 T^2, \quad (10)$$

$$\omega_{1n} = \omega_1 T, \quad (11)$$

$$\omega_{2n} = \omega_2 T, \quad (12)$$

$$\rho = \frac{s(T_0) - \min(s(t))}{\max(s(t)) - \min(s(t))}, \quad (13)$$

$$C_r = \frac{\bar{p} - \min(s(t))}{\max(s(t)) - \min(s(t))}, \quad (14)$$

$$E_r = \frac{\sqrt{a_1^2 + b_1^2}}{\sqrt{a_2^2 + b_2^2}}. \quad (15)$$

Furthermore, in (6), *MCSF* is the mean carotid shape factor, *AIx* is the augmentation index, *SSN* is the supra-sternal notch to femoral site length, *RWAT* is the reflected wave arrival time for a cardiac arterial waveform cycle, *Age* is the age of the participant at the time of tonometry reading,  $P_s$  is the brachial systolic pressure, and  $P_d$  is the brachial diastolic pressure.

In order to provide the most useful subset of these variables into the *PWV* model, we applied a combination of best subset variable selection methods based on multi-linear regression<sup>32</sup>. Using this approach, we ended up with the variables set

$$V = \{\bar{\omega}_2, \omega_{1c}, \omega_{1n}, \rho, E_r, \text{MCSF}, \text{AIx}, \text{SSN}, \text{RWAT}, \text{Age}, P_s, P_d\}. \quad (16)$$

After this stage, we kept 20% of the data as a hold back test set ( $N = 1004$ ). The other part was kept as a train set ( $N = 4016$ ). These sets were picked at random from the original data. However, we chose them in a uniform way such that, both in train and test sets, the *PWV* distribution would follow the population distribution.

After this stage, on the train set, we performed a bootstrap aggregation (bagging) without replacement (sub-sampling)<sup>33</sup> having single layer neural networks at the base regressors with a total number of  $|V|$  neurons

Variable	Units	Mean	Standard Deviation
$\bar{\omega}_2$	—	1.15	0.21
$\omega_{1c}$	$\frac{1}{\text{.}\beta\text{sec}}$	51.46	2.13
$\omega_{1n}$	—	1.45	0.18
$\rho$	—	0.58	0.07
$E_r$	—	2.83	0.65
MCSF	—	0.43	0.03
AIx	—	8.90	13.78
SSN	mm	536.42	39.76
RWAT	msec	134.84	25.62
Age	year	45.15	11.48
$P_s$	mmHg	120.50	14.63
$P_d$	mmHg	68.23	9.85
PWV	m/sec	7.63	2.01

**Table 1.** Population Demographics ( $N = 5020$ ).

in each network. Here,  $|V|$  is the number of elements in  $V$ . The bagging was conducted with sampling 66% at each iteration. A total number of 1000 iteration was used. It is needless to mention that one could stop the iterations when the out-of-bag RMSE reaches a plateau. At each iteration, the neural networks were trained for 100 epochs. A squared penalty of 0.01 was used to prevent over-fitting at each iteration. MATLAB implemented Levenberg-Marquardt backpropagation was used in training the neural nets<sup>34–36</sup>.

The whole machine learning pipeline can be expressed as follows:

1. The uncalibrated waveforms were analysed to extract cardiac cycles;
2. *IF* parameters and waveform features, such as shape factors, were extracted from the selected cardiac cycles;
3. The waveform features were blended in with the routine clinical parameters to construct the original set of features  $V_0$ ;
4. The best subset variable selection method was applied to reduce the dimensionality of the feature-space, namely  $V$ ;
5. A sub-sampled bagged system of neural networks was trained and tested.

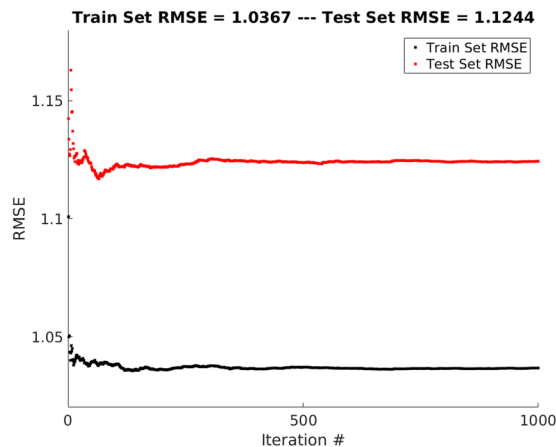
To further analyze our method and the effectiveness of estimated *PWV*, we conducted a prospective cohort study and used proportional hazards regression models to evaluate the association between *PWV* and incident CVD. We evaluated this relationship for *PWV* measurements as well as for estimated *PWV* values produced by our noninvasive *IF* method. Subsequently, we compared the hazards of *PWV* for CVD with that of estimated *PWV*. Baseline population consisted of participants free from CVDs. Adjusted models included components from Framingham risk score: sex, age, total cholesterol, HDL cholesterol, blood pressure, diabetes and smoking. Smoking was defined as regular usage within the last 12-months prior to the examination date. The assumption of proportionality was met. All continuous variables were log transformed to address skewness. Predictive value was evaluated via likelihood ratio test and the Akaike Information Criterion (AIC). Only complete cases without missing data were studied. Kaplan-Meier plots of cumulative probability of a first major CVD event were constructed for *PWV* and also for estimated *PWV*, when participants were grouped according to tertiles of *PWV* and tertiles of estimated *PWV*. Log rank test was used to compare the unadjusted Kaplan-Meier curves. *p*-values < 0.05 were considered as significant.

## Results and Discussion

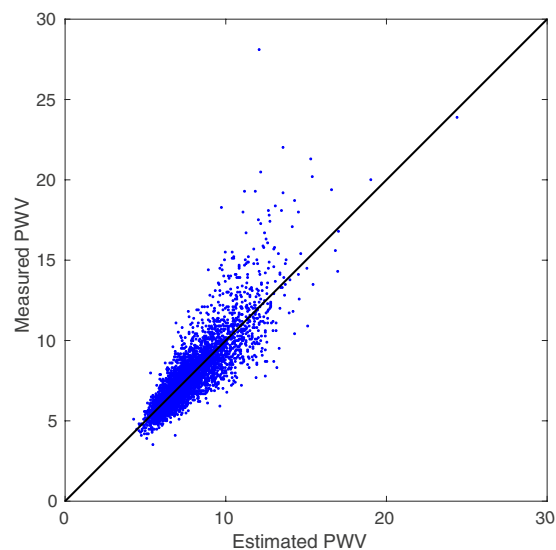
**PWV Model Results.** The population demographics of these variables are shown in Table 1. Convergence of the ensemble of the models was guaranteed by a flat RMSE plot, of both train (RMSE = 1.04) and test (RMSE = 1.12) sets, over the total number of iterations, Fig. 1. The 0.09 gap between the train and test sets, in Fig. 1, shows that the model has an acceptable generalization capability. The total RMSE, on the whole dataset including the train and test sets, was 1.05 m/sec. Our simulations with Decision Tree (DT), boosted DT, boosted Neural Networks<sup>33</sup> show similar but marginally larger RMSE values.

The estimated *PWV* versus the measured values are plotted in Fig. 2. Our model's correlation of the estimated *PWV* with respect to FHS sequential measured *PWV* is 0.85. The Bland-Altman plot of the results is shown in Fig. 3. The limits of agreement are approximately  $\pm 2.07$ .

**Prognosis Results.** Study exclusion criteria resulted in a sample of 4798 usable observations, which included individuals 19 to 69 years old without CVD at the baseline examination. The characteristics of the study sample are presented in Table 2. Within a follow up period of 10 years, 171 participants had a CVD event. Cox proportional hazards models for *PWV* and estimated *PWV* are presented in Table 3. After adjusting for standard risk factors, both *PWV* and estimated *PWV* were significantly associated with an increased risk for a first major CVD



**Figure 1.** RMSE vs Iteration. The RMSE unit is  $m/sec$ . A total number of 1000 iteration is used to make sure that the ensemble of neural networks has converged. The  $0.09 m/sec$  gap between the train and test sets shows that the model has an acceptable generalization capability.



**Figure 2.** Estimated vs Measured PWV: The correlation is 0.85. Both axes have units of  $m/sec$ . The total RMSE is  $1.05 m/sec$ . We can observe that as PWV goes above  $12 m/sec$  the error starts to increase. For smaller values this behavior is less pronounced.

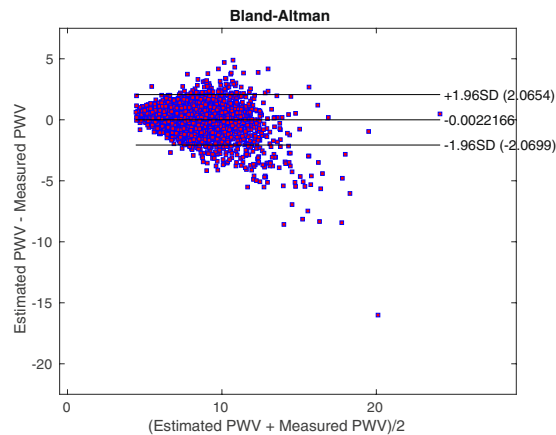
event. Rounding for two digits after zero, the Hazard Ratio (HR) PWV is 4.00, with  $p = 0.0002$ , and the HR of the estimated PWV is 4.43, with  $p = 0.008$ .

**Model Error Analysis.** We can observe, from Figs 2 and 3, that as PWV goes above  $12 m/sec$  the error starts to increase. It is both because of the inherent measurement error<sup>37–40</sup> and also sparsity of data in that region. For smaller values of PWV, this behavior is less pronounced as data concentration is higher and also the measurement errors are smaller.

We can estimate the error of our model compared to the reference method<sup>25</sup> of simultaneous tonometry measurement of PWV. Assign the presented model output with the random variable  $PWV_{predicted}$ . Also, one can label the FHS measured PWV values with the random variable  $PWV_{measured}$ . On the other hand we can name the reference values as  $PWV_{ref}$ . Based on the results of our model, the error of our model with respect to the sequential measurement of PWV is  $1.05 m/sec$ . In other words,

$$std(PWV_{predicted} - PWV_{measured}) = 1.05, \quad (17)$$

where  $std()$  is the standard deviation operator. From the model results, we also have



**Figure 3.** PWV Bland-Altman. Both axes have units of *m/sec*. The limits of agreement are approximately  $\pm 2.07$ . We can again observe that larger values of PWV correspond to more error.

Characteristics	Units	Mean	Standard Deviation
Age	year	44.79	11.32
Sex, Men	%	46	—
$P_s$	mmHg	120.34	14.53
$P_d$	mmHg	68.1	9.8
Pulse Pressure	mmHg	52.25	11.18
HR	bpm	62.4	9.85
BMI	(kg)/(m <sup>2</sup> )	26.35	4.64
Total Cholesterol	(mg)/(dL)	192.54	35.56
HDL Cholesterol	(mg)/(dL)	55.47	16.49
Current smoker	%	16.3	—
Diabetes mellitus	%	3.3	—
PWV	m/sec	7.56	1.93
Estimated PWV	m/sec	7.57	1.64

**Table 2.** Baseline characteristics of the study population ( $N = 4798$ ).

Hemodynamic Measures	Hazard Ratio Including Standard Risk Factors	p-value
PWV	3.998	0.0002
Estimated PWV	4.43	0.008

**Table 3.** PWV and estimated PWV as predictors of a major CVD event ( $N = 4798$ ). Multi-variable model adjusted for age, sex, total cholesterol, HDL cholesterol, pulse pressure, current smoking and diabetes mellitus. All continuous variables were log transformed. This table suggests that both measured PWV and our estimated PWV convey comparable risks for incident CVD in a model adjusted for standard risk factors. These results demonstrate that our estimated PWV is as effective as the measured PWV, such as sequential method, in predicting the risk for CVD.

$$\mathbb{E}(PWV_{predicted} - PWV_{measured}) \approx 0, \tag{18}$$

where  $\mathbb{E}()$  is the expected value operator. Assuming that the error of the sequential PWV measurement is equal to the PulsePen device<sup>41,42</sup>, explained in<sup>25</sup>, we can state

$$std(PWV_{ref} - PWV_{measured}) = 0.31, \tag{19}$$

$$\mathbb{E}(PWV_{ref} - PWV_{measured}) = -0.15. \tag{20}$$



It is logical to state that the error between the reference and sequential methods is independent from the error between our model and sequential methods. As a result, using this fact and Equations (17)–(20), we can deduce that

$$\begin{aligned} & \text{std}(PWV_{\text{predicted}} - PWV_{\text{ref}}) \\ &= \sqrt{\text{std}^2(PWV_{\text{predicted}} - PWV_{\text{measured}}) + \text{std}^2(PWV_{\text{ref}} - PWV_{\text{measured}})} = 1.10, \end{aligned} \quad (21)$$

$$\begin{aligned} & \mathbb{E}(PWV_{\text{predicted}} - PWV_{\text{ref}}) \\ &= \mathbb{E}(PWV_{\text{predicted}} - PWV_{\text{measured}}) + \mathbb{E}(PWV_{\text{measured}} - PWV_{\text{ref}}) = 0.15. \end{aligned} \quad (22)$$

In other words, the estimation error, using the model presented in this paper, is approximately 1.1 *m/sec*. This analysis shows that our model is a reliable surrogate of *PWV*<sup>25</sup>.

**Comparison to other methods.** Our model generated results similar to other non-invasive devices and/or methods currently in use, such as Complior<sup>43</sup>, PulseTrace<sup>44</sup> and Oscillometric<sup>45</sup>.

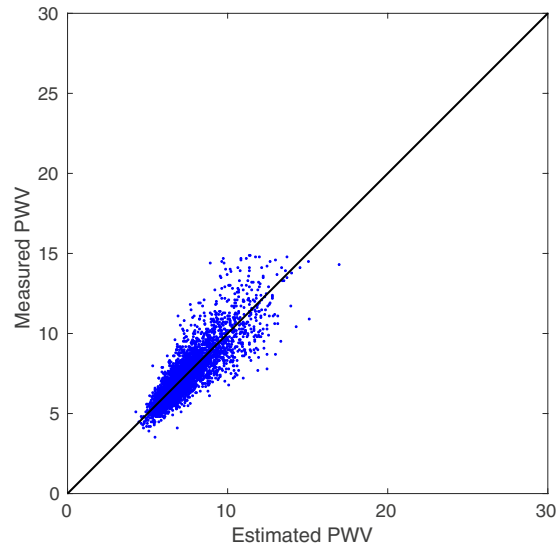
With Complior, there is simultaneous measurement of the pressure pulse. The technician places one probe at the patient's carotid location and a second probe at the femoral location. Then the distance between these two locations is calculated and entered into Complior software. Cuff blood pressure is also measured and entered into the software. *PWV* measures are subsequently generated after a proprietary algorithm is used to measure the pulse transit time between the two locations. With PulseTrace, the stiffness index is estimated by analyzing the photoplethysmographic waves obtained on the fingertip of the individual. The index is calculated by dividing the height of the participant by the time delay between the first systolic peak and the early diastolic peak of the signal. In a study led by Salvi *et al.* in 2008<sup>25</sup>, on a population of 50 participants (aged between 20–84) free from cardiac arrhythmia, these devices produced the following outputs when compared to the reference tonometry method in which waveforms were simultaneously acquired: Complior *Bias* = 2.09 *m/sec* and *LoA* = ±2.68 *m/sec*, *r* = 0.83 and PulseTrace *Bias* = −1.2 *m/sec*, *LoA* = ±4.92 *m/sec*, *r* = 0.55. Although these devices are non-invasive, they are relatively expensive and the procedure is time consuming.

In a prospectively designed validation study led by Feistritz *et al.*<sup>46</sup>, aortic *PWV* estimates from oscillometric technique were generated from 40 participants free of CVDs between 24 and 55 years old. These measurements were compared to the aortic *PWV* values produced by the reference method of cardiac magnetic resonance. Analysis of agreement between the two methods showed *Bias* = 0.57 *m/sec*, *LoA* = ±1.92 *m/sec* and *r* = 0.86. In the aforementioned study, only 28% of the participants were females and the median age of the cohort was 34 years. According to the authors of the paper, the study population does not meet the Artery Society guideline's for *PWV* validation. In specific, the Artery Society needs a homogeneous sex distribution (a minimum of 40% for either sex) as well as a homogeneous distribution along the age groups. Another study designed by Hametner *et al.*<sup>45</sup>, compared oscillometric estimations of aortic *PWV* against intra-aortic arterial *PWV* measurements using a population of 120 patients undergoing elective cardiac catheterization for suspected coronary artery disease (22 patients with *age* ≤ 50 years old and 29 patients with *age* ≥ 70 years old). Exclusion criteria consisted of unstable clinical conditions, arrhythmias and valvular heart disease. In their work, to estimate *PWV* a number of variables from pulse wave analysis and wave separation were combined in a mathematical model in which the major determinants were age, central pressure and aortic characteristic impedance. The following results were then reported by the authors: *Bias* = 0.43 *m/sec*, *LoA* = ±2.45 *m/sec* with a correlation of *r* = 0.81. This study also does not follow the Artery Society guidelines as only 10% of participants were females.

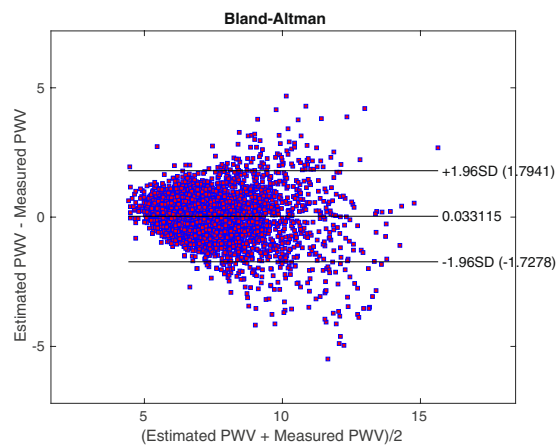
In a recent study by Campo *et al.*<sup>47</sup>, it is shown that the aortic *PWV* can be measured non-invasively with a bathroom scale. The authors combined the principles of ballistocardiography and impedance plethysmography on a single foot to estimate the aortic *PWV*. They compared their *PWV* estimations to measured *PWV*s, on a group of 205 participants. On the validation set, they reported *r* = 0.7, *Bias* = 0.25 *m/sec*, and *LoA* = [−2.48, 2.98]. According to the authors, this new technique presents a few limitations including the gait instability, which affects more frail elderly and some neurological diseases. Other types of diseases may also influence the applicability of the measurement, like for example atrial fibrillation or skin diseases. The population study had several exclusions too. For example, pregnant participants or participants that had morbid obesity (*BMI* > 35) were excluded.

In another recent publication<sup>48</sup>, Greve *et al.* argued that the need for a non-invasive *PWV* estimation is imminent because of the relative inaccessibility of devices such as high-quality applanation tonometry. They have proposed using an equation based on age and mean arterial pressure to perform the estimation. They show that in a healthy group without cardiovascular risk factors, the correlation of the estimation with measured values is *r* = 0.52. This could be seen as a major limitation of that study. Furthermore, they showed that in an apparently healthy patients with cardiovascular risk factors the correlation is *r* = 0.67. Finally, within the group with known CVDs, the correlations is reported to be very low (*r* = 0.37). They also concluded that the estimated *PWV* could predict cardiovascular events independently of the traditional cardiovascular risk factors. However, in another smaller study<sup>49</sup>, Greve *et al.* claimed that the estimated *PWV* only predict CVDs in apparently healthy individuals. Moreover, the estimated *PWV* reclassifies apparently healthy participants to a higher risk category.

**Innovative aspects of proposed method.** The above results show that the estimated *PWV* by Intrinsic Frequencies has the potential to become a reliable non-invasive method of *PWV* measurement. Compared to



**Figure 4.** Estimated vs Measured PWV for  $BMI \leq 30$  and  $PWV < 15$ . Both axes have units of  $m/sec$ . Using the Artery Society guidelines for validation of non-invasive hemodynamic measurement devices, we observe that there is a strong agreement between our estimation of PWV and the recorded values.



**Figure 5.** PWV Bland-Altman for  $BMI \leq 30$  and  $PWV < 15$ . Both axes have units of  $m/sec$ . This plot shows that our PWV estimation would be graded as acceptable by the Artery Society.

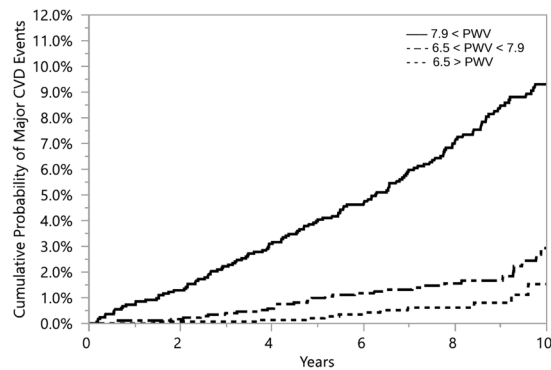
Complior, PulseTrace and Oscillometric, the IF technique is less intrusive and easier to operate that offers a more practical solution to PWV estimations.

The presented study uses an artificial intelligence technique to render an accurate estimation of central arterial stiffness (carotid-femoral PWV). The sample size used in this study is large enough and includes both healthy and CVD volunteers to offer an adequate statistical power. This study also uses a homogeneous age and gender distribution population.

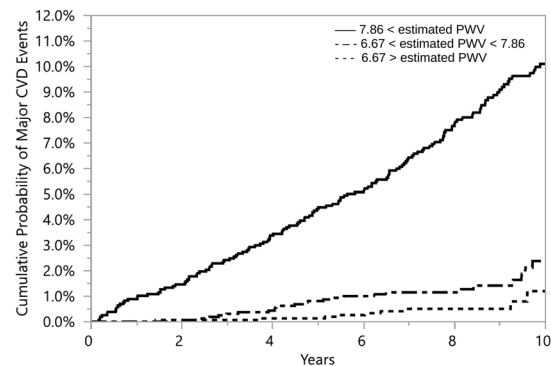
We further analyzed our results based on the Artery Society guidelines for validation of non-invasive hemodynamic measurement devices<sup>50</sup>. We subsequently excluded the following from our results before generating updated Bland-Altman plots: individuals with a  $BMI \geq 30$  (due to problems regarding the measurement of an accurate path length) and  $PWV \geq 15$ . The updated results after the fore-mentioned filtering were:  $Bias = 0.03 m/sec$ ,  $LoA = \pm 1.76 m/sec$  ( $SD = 0.88$ ); see Figs 4 and 5. These outputs would be graded as acceptable by the Artery Society.

**Risk Evaluation.** The results, see Table 3, suggest that PWV and our estimated PWV both convey comparable risks for incident CVD in a model adjusted for standard risk factors. The use of Kaplan-Meier failure method showed that when participants were grouped according to tertiles of PWV, the probability of developing a CVD event increased with the group displaying higher PWV values (log-rank test  $p < 0.0001$ ); see Fig. 6. The same observation was made when participants were grouped by tertiles of the estimated PWV. The group with estimated PWV of  $7.86 m/sec$  or higher was at an increased risk of developing a CVD event (log-rank test  $p < 0.0001$ );





**Figure 6.** Kaplan–Meier estimates of cardiovascular disease by tertiles of  $PWV$  index. The solid line represents  $7.9 \leq PWV$ . The solid dashed line represents  $6.5 \leq PWV < 7.9$ . The dashed line represents  $PWV \leq 6.5$ .



**Figure 7.** Kaplan–Meier estimates of cardiovascular disease by tertiles of estimated  $PWV$  index. The solid line represents  $7.86 \leq \text{estimated } PWV$ . The solid dashed line represents  $6.67 \leq \text{estimated } PWV < 7.86$ . The dashed line represents  $\text{estimated } PWV \leq 6.67$ .

see Fig. 7. These results show that estimated  $PWV$  by Intrinsic frequencies is as effective as  $PWV$  measurements obtained by direct methods, such as sequential method, in predicting the risk for CVD.

**Model Performance.** Following the Kaplan-Meier failure method analysis, we segmented the  $PWV$  data into three different groups, to check the model performance for different subsets according to  $PWV$  values. Figure 8 shows that in all three tertiles  $PWV \leq 6.5$ ,  $6.5 \leq PWV < 7.9$  and  $PWV \geq 7.9$ , the model presented in this paper has an acceptable performance.

**Study Limitations.** The major limitations of this study are:

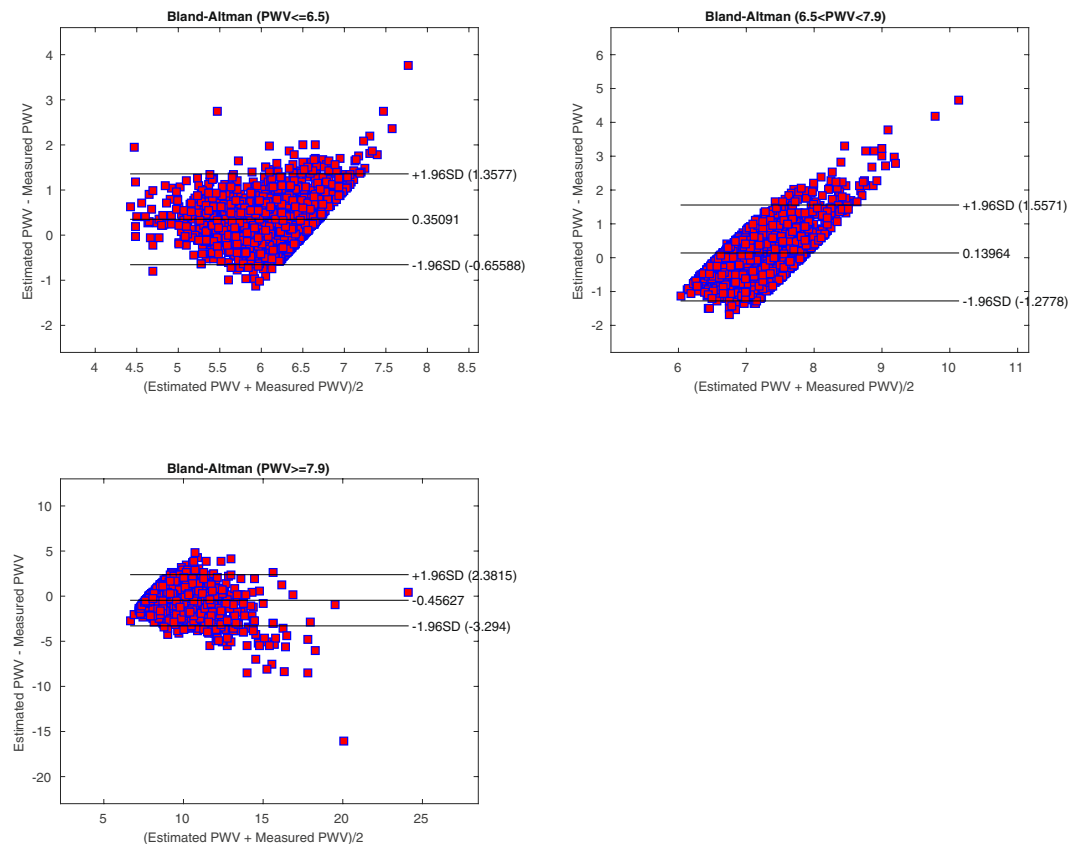
1. The population used in this study was not racially diverse (mostly Caucasian). For a more general conclusion, we need to have a more diverse study population.
2. The automatic cycle selection, used in this study, is prone to mis-identification of cardiac cycles and diastolic notches. However, the effect of the related error on overall finding of this study is insignificant.
3. The sequential method used to calculate the  $PWV$  might depict error either in body surface measurements of wave travel times at high  $PWV$  values.

## Conclusions

In this paper, we have introduced a novel artificial intelligence method to estimate the carotid-femoral pulse wave velocity. This method is based on the newly introduced Intrinsic Frequency method<sup>19</sup> and as inputs uses only an uncalibrated carotid pressure waveform with typical clinical variables such as blood pressure.

The main advantages in having an estimated  $PWV$  from an uncalibrated carotid pressure waveform, with few typical clinical variables such as blood pressure, would be that it does not need an ECG measurement nor a femoral tonometry recording. As a result, it is easier, and potentially can be done by a smart phone as we have shown in our previous publication that carotid waveform can be easily measured using a regular iPhone camera<sup>20</sup>.

Here, in this article, we have been able to address the need of estimating  $PWV$  by providing an accurate and precise statistical model estimating pulse wave velocity. The model presented in this manuscript can estimate  $PWV$  with an RMSE of 1.12 m/sec, compared to the reference method.



**Figure 8.** PWV model Bland-Altman for tertiles. All axes have units of  $m/sec$ . The tertiles are coming from the Kaplan-Meier analysis of the original PWV values. The upper-left plot is the Bland-Altman for  $PWV \leq 6.5$ . The upper-right plot is for  $6.5 < PWV < 7.9$ . The lower-left plot is the Bland-Altman for  $7.9 \leq PWV$ . All three different tertiles show that the model presented in this paper have an acceptable performance.

We have provided an error analysis and comparison to other methods currently in use in order to support the conclusion that the presented model is an acceptable surrogate for arterial stiffness. Furthermore we conducted a prospective investigation to analyze the predictive value of estimated PWV in relation to the onset of CVDs. Our study showed that estimated PWV was significantly associated with increased risk of CVDs.

**Data Availability.** The data used in this study can be requested from FHS directly. It is publicly available to qualified investigators. An approved research proposals could be qualified to receive the de-identified data. FHS data, in general, can be requested by a research application submission to one of the following:

- Directly from Framingham Heart Study (<https://www.framinghamheartstudy.org/>), BioLINCC (<https://biolincc.nhlbi.nih.gov/home/>), or
- dbGaP (<https://www.ncbi.nlm.nih.gov/gap>).

The manuscript data can be found using the following links:

1. (<https://biolincc.nhlbi.nih.gov/studies/gen3/?q=framingham>) - Gen3 cohort
2. (<https://biolincc.nhlbi.nih.gov/studies/framcohort/?=framingham>) - Original Cohort
3. (<https://biolincc.nhlbi.nih.gov/studies/framoffspring/?q=framingham>) - Offspring Cohort.

**Research Ethics.** All experiments and procedures were reviewed and approved by FHS. The Framingham Heart Study is conducted and supported by the National Heart Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01- HC-25195).

**Permission to Carry Out Fieldwork.** This study did not have fieldwork.

## References

1. Lloyd-Jones, D. *et al.* Heart disease and stroke statistics 2010 update. *Circulation* **121**, e46–e215 (2010).
2. Mozaffarian, D. *et al.* Executive summary: Heart disease and stroke statistics-2016 update: A report from the american heart association. *Circulation* **133**, 447 (2016).

3. Laurent, S. *et al.* Aortic stiffness is an independent predictor of all-cause and cardiovascular mortality in hypertensive patients. *Hypertension* **37**, 1236–1241 (2001).
4. Sutton-Tyrrell, K. *et al.* Elevated aortic pulse wave velocity, a marker of arterial stiffness, predicts cardiovascular events in well-functioning older adults. *Circulation* **111**, 3384–3390 (2005).
5. Laurent, S. *et al.* Expert consensus document on arterial stiffness: methodological issues and clinical applications. *European heart journal* **27**, 2588–2605 (2006).
6. Mitchell, G. F. *et al.* Changes in arterial stiffness and wave reflection with advancing age in healthy men and women the framingham heart study. *Hypertension* **43**, 1239–1245 (2004).
7. Mitchell, G. F. *et al.* Arterial stiffness and cardiovascular events the framingham heart study. *Circulation* **121**, 505–511 (2010).
8. Townsend, R. R. Arterial stiffness: recommendations and standardization. *Pulse* **4**, 3–7 (2016).
9. Blacher, J. *et al.* Impact of aortic stiffness on survival in end-stage renal disease. *Circulation* **99**, 2434–2439 (1999).
10. Choi, C. U. *et al.* Impact of aortic stiffness on cardiovascular disease in patients with chest pain: assessment with direct intra-arterial measurement. *American journal of hypertension* **20**, 1163–1169 (2007).
11. Cruickshank, K. *et al.* Aortic pulse-wave velocity and its relationship to mortality in diabetes and glucose intolerance. *Circulation* **106**, 2085–2090 (2002).
12. Hansen, T. W. *et al.* Prognostic value of aortic pulse wave velocity as index of arterial stiffness in the general population. *Circulation* **113**, 664–670 (2006).
13. Mattace-Raso, F. U. *et al.* Arterial stiffness and risk of coronary heart disease and stroke. *Circulation* **113**, 657–663 (2006).
14. Shokawa, T. *et al.* Pulse wave velocity predicts cardiovascular mortality. *Circulation Journal* **69**, 259–264 (2005).
15. Rajzer, M. W. *et al.* Comparison of aortic pulse wave velocity measured by three techniques: Complior, sphygmocor and arteriograph. *Journal of hypertension* **26**, 2001–2007 (2008).
16. Pereira, T., Correia, C. & Cardoso, J. Novel methods for pulse wave velocity measurement. *Journal of medical and biological engineering* **35**, 555–565 (2015).
17. Ramesh, A., Kambhampati, C., Monson, J. & Drew, P. Artificial intelligence in medicine. *Annals of The Royal College of Surgeons of England* **86**, 334 (2004).
18. Pahlevan, N. M. *et al.* Intrinsic frequency for a systems approach to haemodynamic waveform analysis with clinical applications. *Journal of The Royal Society Interface* **11**, 20140617 (2014).
19. Tavallali, P., Hou, T. Y., Rinderknecht, D. G. & Pahlevan, N. M. On the convergence and accuracy of the cardiovascular intrinsic frequency method. *Royal Society Open Science* **2**, 150475 (2015).
20. Pahlevan, N. M. *et al.* Noninvasive iphone measurement of left ventricular ejection fraction using intrinsic frequency methodology. *Critical care medicine* (2017).
21. Framingham heart study. <https://www.framinghamheartstudy.org/>. Accessed: (2016).
22. Splansky, G. L. *et al.* The third generation cohort of the national heart, lung, and blood institute's framingham heart study: design, recruitment, and initial examination. *American journal of epidemiology* **165**, 1328–1335 (2007).
23. Kannel, W. B., Feinleib, M., McNAMARA, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families the framingham offspring study. *American journal of epidemiology* **110**, 281–290 (1979).
24. Dawber, T. R., Meadors, G. F. & Moore, F. E. Jr. Epidemiological approaches to heart disease: The framingham study\*. *American Journal of Public Health and the Nations Health* **41**, 279–286 (1951).
25. Salvi, P. *et al.* Comparative study of methodologies for pulse wave velocity estimation. *Journal of human hypertension* **22**, 669–677 (2008).
26. O'Rourke, M. F. & Gallagher, D. E. Pulse wave analysis. *Journal of Hypertension-Supplement* **14**, S147–S158 (1996).
27. Fan, Z., Zhang, G. & Liao, S. Pulse wave analysis. In *Advanced Biomedical Engineering* (InTech, 2011).
28. Li, B. N., Dong, M. C. & Vai, M. I. On an automatic delineator for arterial blood pressure waveforms. *Biomedical Signal Processing and Control* **5**, 76–81 (2010).
29. Zong, W., Heldt, T., Moody, G. & Mark, R. An open-source algorithm to detect onset of arterial blood pressure pulses. In *Computers in Cardiology*, **2003**, 259–262 (IEEE, 2003).
30. Tavallali, P., Razavi, M. & Brady, S. A non-linear data mining parameter selection algorithm for continuous variables. *PLoS one* **12**, e0187676 (2017).
31. Determinants of pulse wave velocity in healthy people and in the presence of cardiovascular risk factors: establishing normal and reference values. *European heart journal* **31**, 2338–2350 (2010).
32. Furnival, G. M. & Wilson, R. W. Regressions by leaps and bounds. *Technometrics* **42**, 69–79 (2000).
33. Friedman, J., Hastie, T. & Tibshirani, R. *The elements of statistical learning*, vol. 1 (Springer series in statistics Springer, Berlin, 2001).
34. Demuth, H. B., Beale, M. H., De Jess, O. & Hagan, M. T. *Neural network design* (Martin Hagan, 2014).
35. Hagan, M. T. & Menhaj, M. B. Training feedforward networks with the marquardt algorithm. *IEEE transactions on Neural Networks* **5**, 989–993 (1994).
36. Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics* **11**, 431–441 (1963).
37. Karamanoglu, M. Errors in estimating propagation distances in pulse wave velocity. *Hypertension* **41**, e8–e8 (2003).
38. Mestre, C., Lantelme, P., Lievre, M., Gressard & Milon, H. Heart rate: An important confounder of pulse wave velocity assessment. In *Journal of Hypertension*, vol. 20, S183–S183 (Lippincott Williams & Wilkins 530 Walnut St, Philadelphia, PA 19106-3621 USA, 2002).
39. Segers, P. *et al.* Limitations and pitfalls of non-invasive measurement of arterial pressure wave reflections and pulse wave velocity. *Artery Research* **3**, 79–88 (2009).
40. Weir-McCall, J. R. *et al.* Effects of inaccuracies in arterial path length measurement on differences in mri and tonometry measured pulse wave velocity. *BMC cardiovascular disorders* **17**, 118 (2017).
41. Othmane, T. E. H. *et al.* Effect of sevelamer on aortic pulse wave velocity in patients on hemodialysis: a prospective observational study. *Hemodialysis International* **11**, S13–S21 (2007).
42. Salvi, P. *et al.* Validation of a new non-invasive portable tonometer for determining arterial pressure wave and pulse wave velocity: the pulsepen device. *Journal of hypertension* **22**, 2285–2293 (2004).
43. Complior. <http://www.complior.com/products/>. Accessed: (2017).
44. Pulsetrace. <http://www.micromedical.co.uk/>. Accessed: (2017).
45. Hametner, B. *et al.* Oscillometric estimation of aortic pulse wave velocity: comparison with intra-aortic catheter measurements. *Blood pressure monitoring* **18**, 173–176 (2013).
46. Feistritz, H.-J. *et al.* Comparison of an oscillometric method with cardiac magnetic resonance for the analysis of aortic pulse wave velocity. *PLoS one* **10**, e0116862 (2015).
47. Campo, D. *et al.* Measurement of aortic pulse wave velocity with a connected bathroom scale. *American Journal of Hypertension* **hpx059** (2017).
48. Greve, S. V. *et al.* Estimated carotid–femoral pulse wave velocity has similar predictive value as measured carotid–femoral pulse wave velocity. *Journal of hypertension* **34**, 1279–1289 (2016).
49. Greve, S. V., Laurent, S. & Olsen, M. H. Estimated pulse wave velocity calculated from age and mean arterial blood pressure. *Pulse* **4**, 175–179 (2016).
50. Wilkinson, I. B. *et al.* Artery society guidelines for validation of non-invasive haemodynamic measurement devices: Part 1, arterial pulse wave velocity. *Artery Research* **4**, 34–40 (2010).

## Acknowledgements

The Framingham Heart Study is supported by Contract No. HHSN268201500001I from the National Heart, Lung, and Blood Institute (NHLBI) with additional support from other sources. This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or conclusions of the Framingham Heart Study or the NHLBI.

## Author Contributions

P.T. conceived of the mathematical and numerical methods of the study; carried out the modeling pipeline design; programmed the cycle-notch selection method; designed and programmed the IF algorithm; designed and programmed the ensemble learning code of the method; helped with finding the optimum hyper-parameters of the ensemble learning code; helped with risk analysis; helped with variable selection; and drafted the manuscript. M.R. conducted the variable selection; conducted the literature survey; found the optimum hyper-parameters of the ensemble learning code; conducted risk analysis; and helped draft and revise the manuscript. N.M.P. helped draft and revise the manuscript.

## Additional Information

**Competing Interests:** Dr. Tavallali had employment agreement with Avicena at the time of the manuscript preparation. Dr. Razavi has employment agreement with Avicena. Dr. Pahlevan has consulting agreements with Avicena.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018