



OPEN

## Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure

Eoin Finnegan<sup>1✉</sup>, Shaun Davidson<sup>1</sup>, Mirae Harford<sup>1,2,3</sup>, Peter Watkinson<sup>2,3</sup>, Lionel Tarassenko<sup>1</sup> & Mauricio Villarroel<sup>1</sup>

There is a growing emphasis being placed on the potential for cuffless blood pressure (BP) estimation through modelling of morphological features from the photoplethysmogram (PPG) and electrocardiogram (ECG). However, the appropriate features and models to use remain unclear. We investigated the best features available from the PPG and ECG for BP estimation using both linear and non-linear machine learning models. We conducted a clinical study in which changes in BP ( $\Delta$ BP) were induced by an infusion of phenylephrine in 30 healthy volunteers (53.8% female, 28.0 (9.0) years old). We extracted a large and diverse set of features from both the PPG and the ECG and assessed their individual importance for estimating  $\Delta$ BP through Shapley additive explanation values and a ranking coefficient. We trained, tuned, and evaluated linear (ordinary least squares, OLS) and non-linear (random forest, RF) machine learning models to estimate  $\Delta$ BP in a nested leave-one-subject-out cross-validation framework. We reported the results as correlation coefficient ( $\rho_p$ ), root mean squared error (RMSE), and mean absolute error (MAE). The non-linear RF model significantly ( $p < 0.05$ ) outperformed the linear OLS model using both the PPG and the ECG signals across all performance metrics. Estimating  $\Delta$ SBP using the PPG alone ( $\rho_p = 0.86$  (0.23), RMSE = 5.66 (4.76) mmHg, MAE = 4.86 (4.29) mmHg) performed significantly better than using the ECG alone ( $\rho_p = 0.69$  (0.45), RMSE = 6.79 (4.76) mmHg, MAE = 5.28 (4.57) mmHg), all  $p < 0.001$ . The highest ranking features from the PPG largely modelled increasing reflected wave interference driven by changes in arterial stiffness. This finding was supported by changes observed in the PPG waveform in response to the phenylephrine infusion. However, a large number of features were required for accurate BP estimation, highlighting the high complexity of the problem. We conclude that the PPG alone may be further explored as a potential single source, cuffless, blood pressure estimator. The use of the ECG alone is not justified. Non-linear models may perform better as they are able to incorporate interactions between feature values and demographics. However, demographics may not adequately account for the unique and individualised relationship between the extracted features and BP.

Changes in the cardiovascular and autonomic nervous systems are reflected in changes in signals such as the photoplethysmogram (PPG) and the electrocardiogram (ECG). These physiological signals are ubiquitous in a clinical setting and increasingly in an out-of-clinic setting due to the development of wearables such as smart-watches. As a result, recent advances in non-invasive, cuffless estimation of blood pressure (BP) have been focused on utilising the PPG and the ECG signals. For example, pulse arrival time (PAT), computed as the time difference between two fiducial points in the ECG and PPG waveforms, has been shown in certain studies to have strong correlations with BP<sup>1</sup>. However, measuring PAT requires two synchronous devices, both of which are susceptible to independent sources of noise such as motion artefacts. Additionally, factors such as the pre-ejection period (PEP) (the time delay between the electrical depolarisation of the heart's left ventricle and the

<sup>1</sup>Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK. <sup>2</sup>Critical Care Research Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK. <sup>3</sup>NIHR Oxford Biomedical Research Centre, Oxford, UK. ✉email: eoin.finnegan@eng.ox.ac.uk

opening of the aortic valve) may impact PAT estimates independently of BP<sup>2</sup>. A summary of BP estimation methods using PAT can be found in<sup>3–5</sup>.

Driven by modern techniques, research on cuffless BP estimation has increasingly focused on relating morphological features of the PPG and the ECG waveforms to BP (or changes in BP,  $\Delta$ BP) using data-driven models. However, the optimal features and models required for accurate BP estimation remain unclear. In this paper, we implemented linear and non-linear machine learning (ML) models to estimate  $\Delta$ BP using a large and diverse cohort of features from the PPG and the ECG waveforms. In addition to features commonly used in the literature, we proposed new features from both signals and assessed their individual importance for estimating  $\Delta$ BP using Shapley additive explanation values. We compared our results to those estimating changes in BP using PAT, and evaluated the PPG and ECG as potential single source devices for BP estimation. This work was carried out using data from a clinical study involving 30 healthy volunteers. Changes in BP were induced by the administration of phenylephrine, a vasoactive medication that causes arterial and venous vasoconstriction and increases cardiac preload (initial stretching of cardiac muscles).

**Relationship between changes in BP and changes in the PPG.** The pulsatile PPG waveform is related to changes in blood volume over time in a bed of tissue<sup>6</sup>. The PPG signal is typically recorded by a pulse oximeter placed on the index finger. Methods using the PPG have been proposed as a continuous, non-invasive, cuffless approach to estimate BP<sup>7</sup>. The PPG–BP relationship is driven, in part, by the theoretical relationship between changes in pressure and volume of blood in a localised region of the arteries<sup>7</sup>, as well as the impact of reflected pressure waves<sup>8</sup>. Reflected pressure waves result at points of significant impedance mismatch further down the arterial tree and travel back to the point of PPG measurement. The first reflection site is at the junction of the renal arteries resulting in a reflected pressure wave (known as the tidal wave) that is typically present in late systole<sup>9</sup>. The second reflection site is at the junction of the iliac arteries resulting in a reflected pressure wave (known as the dicrotic wave) that is typically present in early diastole<sup>10</sup>. Changes in arterial stiffness are thought to have a significant impact on the timing, amplitude, and morphology of the reflected pressure waves and therefore may significantly impact the shape of the PPG waveform<sup>8</sup>. The PPG offers significant benefits for BP monitoring over conventional cuff-based measurements. Most notably, the PPG can be recorded by a single, unobtrusive optical sensor which also has the potential to be implemented on a wearable device such as a smartwatch<sup>11,12</sup>. However, there is no generally accepted method relating changes in the PPG waveform to  $\Delta$ BP and, as a result, a variety of different approaches are proposed in the literature. A summary of BP estimation algorithms using the PPG waveform can be found in<sup>7,13,14</sup>.

Much like PAT<sup>1</sup>, features extracted from the PPG waveform are thought to be subject-specific requiring calibration for accurate mapping to BP values<sup>7</sup>. Mukkamala et al.<sup>7</sup> splits calibration strategies into three groups: *individual*, *hybrid* and *population*. In *individual calibration*, all model parameters are determined by multiple paired recordings of BP and PPG from a single individual. This approach may be feasible for BP estimation using PAT (where typically only two parameters are required for modelling<sup>1</sup>), however it becomes intractable for multi-parameter ML models often employed for PPG-based BP estimation. In *hybrid calibration*, only one calibration BP data point is required for a single individual. The remaining model parameters are estimated using the individual's demographics and a training set comprised of multiple BP-PPG pairs from a cohort of different individuals. In *population calibration*, no calibration recordings are required and all calibration is handled using the individual's demographics and a similar training set. Certain dependencies on the morphology of the PPG waveform have been previously reported for age<sup>15</sup>, sex<sup>16</sup> and body mass index (BMI)<sup>17</sup>, however these are often not strong enough to allow for good accuracy when using *population calibration* strategies. As a result, the majority of studies proposing PPG-based BP estimation opt for a *hybrid calibration* strategy.

Sun et al.<sup>18</sup> evaluated the use of a linear regression model for estimating systolic blood pressure (SBP) measured by a commercial Portapres device (Finapres), using the volume-clamp method. Nineteen subjects underwent an exercise test followed by a posture change test. Nineteen features were extracted from the PPG. Using a *hybrid calibration* strategy, the authors reported a root mean squared error (RMSE) of 8.99 mmHg and Pearson's correlation coefficient ( $\rho_p$ ) of 0.85 during the exercise test, and a RMSE of 7.33 mmHg and  $\rho_p$  of 0.47 during the posture change test. Normalised weights of the linear model were used to highlight features that had the most predictive power and an inconsistency of the best features was found between the two tests. Miao et al.<sup>19</sup> used support vector regression on 14 extracted features from 73 subjects to track BP changes induced by physical exercise. A genetic algorithm for feature selection was developed to highlight features that best estimated  $\Delta$ BP. The stability of the proposed models was evaluated in a follow-up study for 1 day, 10 days and 6 months after the initial test. The results suggested that, similar to PAT<sup>20,21</sup>, the models lose their accuracy over time and therefore require frequent recalibration. Hasanzadeh et al.<sup>22</sup> extracted features from the PPG in a subset of 1,000 individuals from the MIMIC-III dataset. The intra-arterial blood pressure was used as a reference. The authors implemented a linear regression model as well as a non-linear tree-based model, AdaBoost. AdaBoost significantly outperformed the linear model for their dataset. Additionally, the authors highlighted the sensitivity of PPG feature detection in the presence of random noise. Recent shifts in the field have moved towards the application of deep learning algorithms<sup>23,24</sup> which often do not require the extraction of handcrafted features and instead work on the raw PPG waveform. However, while these approaches are reported to improve the estimation accuracy, they lose out in model interpretability due to their black-box nature.

**Relationship between changes in BP and changes in the ECG.** The ECG is a measure of the electrical activity generated in the myocardium (heart muscle) during each heartbeat. It is acquired by measuring the voltage difference between two points on the body surface over time<sup>25</sup>. A single lead ECG is typically recorded by three electrodes placed on an individual's torso forming Einthoven's triangle<sup>25</sup>.

In comparison to the PPG, less focus has been placed on the potential use of the ECG for BP estimation. The general theory governing the relationship between ECG and BP is based on a cyclical process known as mechano-electric coupling (MEC)<sup>26</sup>. Changes in the electrical properties of the heart have a direct impact on its contractility. This is known as excitation-contraction coupling. Similarly, changes in the mechanical properties of the tissues surrounding the heart are detected by mechanosensitive ion channels, resulting in local changes in the electrical potential. This is known as mechano-electrical feedback. MEC, therefore, describes the cyclical process whereby changes in the ECG waveform can reflect  $\Delta$ BP. However, MEC is influenced not only by extra-cardiac control mechanisms such as the Autonomic Nervous System (ANS) and hormonal changes, but also by environmental mechanisms such as ion concentrations and temperature<sup>27</sup>. Therefore, the relationship between  $\Delta$ BP that can be detected by analysing morphological changes in the ECG waveform are not as developed as methods using the PPG waveform.

Simjanoska et al.<sup>28</sup> used a dataset containing 51 individuals from a mixture of 4 open-source datasets. Three of the datasets (totalling 44 individuals) included healthy volunteers using commercial ECG sensors with reference BP values measured by a cuff. In these datasets BP was perturbed by natural variations with each individual contributing a range from 1 to 8 measurements overall. The fourth dataset was recorded from 7 patients with traumatic brain injuries in an intensive care unit (ICU) using clinical-grade ECG sensors and reference BP values measured by an arterial line. Complexity features such as mobility and entropy were extracted from 30-second segments of ECG data and used to train a random forest regression model using a train-validation-test split. After calibration, the model achieved mean absolute error (MAE) of 7.72 mmHg for SBP, 9.45 mmHg for diastolic blood pressure (DBP), and 8.13 mmHg for mean arterial blood pressure (MAP). However, the authors only implemented a small number of complexity features, and the mixture of data sources (especially with regard to mixing healthy and unhealthy populations) makes interpreting the results difficult.

## Methods

**Clinical study.** Thirty healthy volunteers with no history of cardiovascular disease were recruited for the clinical study. The study protocol has been outlined previously<sup>1,29</sup>. The study took place at the Cardiovascular Clinical Research Facility within the John Radcliffe Hospital, Oxford, UK. This study was reviewed and approved by the Oxford University Research and Ethics Committee and Clinical Trials and Research Governance teams (R63796/RE001). All methods were performed in accordance with the relevant guidelines and regulations. Individual informed written consent was obtained from all the participants in the study to record the data and publish the results.

The study was split into four main stages: rest, dose increase, maximum infusion, and washout. Each session began with a 5-minute resting period. Participants were then administered an infusion of phenylephrine. Phenylephrine is an  $\alpha_1$  adrenergic receptor agonist which causes arterial and venous vasoconstriction as well as reflex bradycardia (slowing down of heart rate)<sup>30</sup>. The dosing regimen was guided by the clinician in the Medical Sciences Division Ethics Committee (University of Oxford) who balanced the desired clinical effect against any safety concerns<sup>29</sup>. Phenylephrine was administered as an intravenous solution, starting at a rate of 0.2mcg/kg/min with an increase of 0.2mcg/kg/min every 1 minute for 10 increments. Once the maximum rate was reached, the infusion and all monitoring remained constant for a further 6 minutes. Each session ended with an 8-minute washout period. Participants were asked to refrain from ingesting caffeinated drinks for four hours prior to the study visit as caffeine is a vasoconstrictor. Participants lay with the head and trunk raised to between 15 and 45 degrees (semi-Fowler's position).

**Instrumentation.** A Philips Intellivue MX800 patient monitor (Philips, Netherlands) recorded BP via a sphygmomanometer cuff. The Philips comfort care M3001A BP cuff was wrapped around the upper left arm of the participant with the centre directly above the brachial artery as recommended<sup>31</sup>. The ixTrend software (Ixellence GmbH, Germany) was used to record the data generated by the patient monitor. A Stowood Visi Black shadow polygraphy device (Stowood, UK) recorded the ECG and PPG signals. Three ECG electrodes were attached to the participant's torso forming Einthoven's triangle. A pulse oximeter (Masimo, USA) probe was placed on the participant's right hand. Table 1 provides a summary of the physiological parameters recorded by the Philips and Stowood monitors.

**Features from the PPG.** To preprocess the PPG waveform, we performed the following steps. The PPG signal was first filtered using an 8<sup>th</sup>-order Butterworth infinite impulse response (IIR) band-pass filter with cut-off frequencies of 0.5 Hz and 10 Hz as recommended in<sup>3</sup>. We then followed the work of Villarroel et al.<sup>32</sup> to detect the pulse onsets and to assess the quality of the PPG signal ( $SQI_{PPG}$ ). Amplitude modulation of the PPG can result from respiration<sup>33</sup> as well as changes in contact pressure of the sensor on the surface of the skin<sup>34</sup>. In order to reduce the dependence of this amplitude modulation as well as heart rate, we normalised the PPG beats

Device	Name	Description	Sampling rate (Hz)
Philips	SBP, MAP, DBP	Blood pressure signal from the sphygmomanometer cuff	1/60
Stowood	ECG	1-lead electrocardiography signal	256
	PPG	Photoplethysmography signal	512
	HR	Heart rate derived from Masimo pulse oximeter	5

**Table 1.** Physiological parameters and sampling rates recorded by the two devices used in this study.

to have unit amplitude and time duration. The first and second derivatives of the PPG beats, known as the velocity plethysmogram (VPG) and acceleration plethysmogram (APG) respectively, were extracted using a 7<sup>th</sup>-order Savitzky-Golay filter<sup>35</sup>. We used the nomenclature defined by Elgendi et al.<sup>36</sup> for the characteristic fiducial points of the PPG, VPG, and APG.

Figure 1 shows the typical changes in the morphology of the PPG, VPG and APG beat seen for a participant during the four main stages of the study protocol (rest, dose increase, maximum infusion and washout). The following fiducial points were detected using the criteria defined in<sup>37</sup>: dicotic notch (N), diastolic peak (D) and the *a*, *b*, *c*, *d*, and *e* waves on the APG. The dicotic notch is produced by the closing of the aortic valve and therefore marks the end of systole and the beginning of diastole. The systolic peak (S) was defined as the first turning point of the PPG pulse above its midpoint.

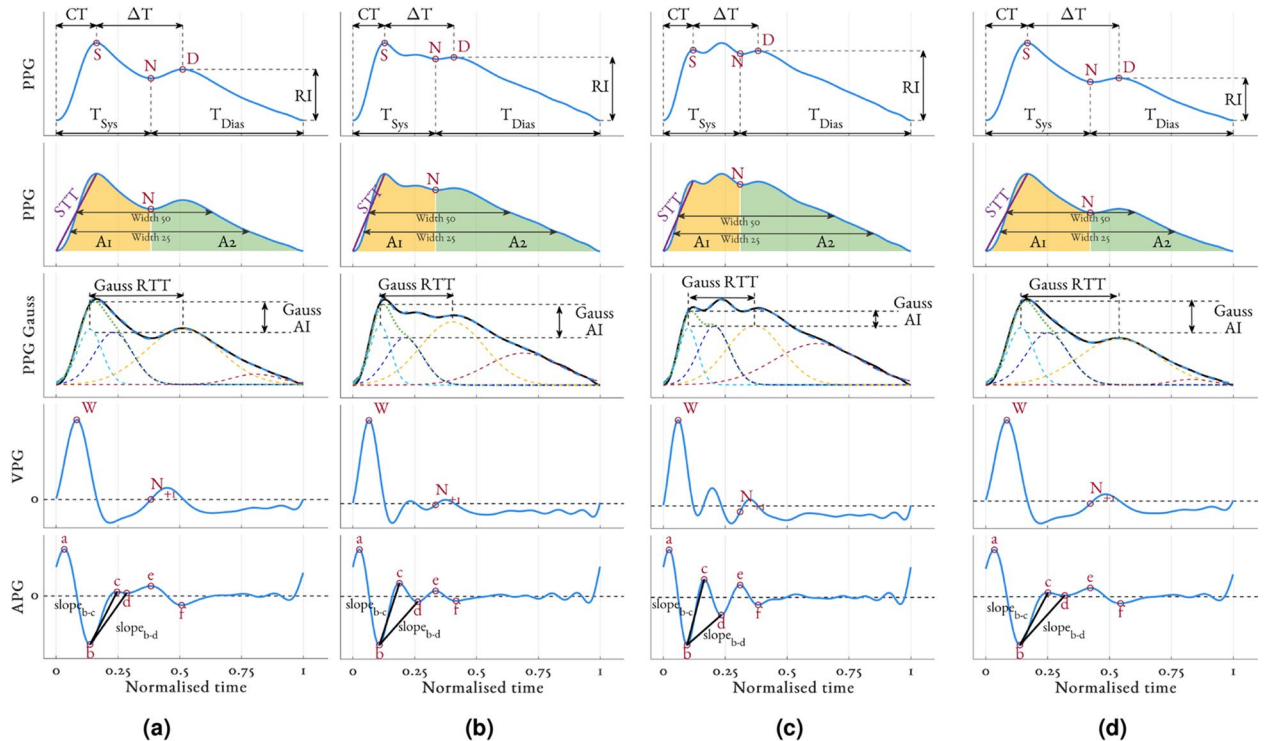
We identified features from previous publications describing methods for estimating  $\Delta$ BP from the PPG. Table 2 provides a summary of the features we extracted from the PPG. We grouped these features into five main categories: PPG morphology, VPG morphology, APG morphology, Gaussian decomposition, and principal component analysis (PCA).

**PPG morphology.** As BP is determined by the product of total peripheral resistance (TPR) and cardiac output (CO), features reflecting variations in these two metrics are likely to provide a strong indication of  $\Delta$ BP.

Changes in TPR can significantly impact the morphology of the PPG through interactions with the incident forward travelling wave produced by left ventricular contraction and the reflected pressure waves<sup>8</sup>. To model the varying amplitude of the reflected pressure waves, we included the dicotic notch amplitude ( $N_{amp}$ ), and the diastolic peak amplitude (labelled reflection index, RI, for consistent feature notation with the literature) as features<sup>38</sup>. As the tidal and dicotic reflected waves typically arrive during systole and diastole respectively, Lyu et al.<sup>39</sup> proposed the stress-induced vascular response index (sVRI) as the ratio of the mean of the PPG in the diastolic phase,  $\mu_{Dias}$ , to the mean of the PPG in the systolic phase,  $\mu_{Sys}$ . This effect was additionally modelled using the areas under PPG during the systolic (A1) and diastolic (A2) phases<sup>40</sup>. The ratio of these two areas (A2/A1) is referred to as the inflection point area ratio (IPA).

The influence of the reflected pressure waves was additionally modelled through analysis of the frequency domain by assessing the level of distortion in the PPG waveform as the relative power of its harmonic frequencies using the normalised harmonic area (NHA)<sup>40</sup>.

The increased velocity of the reflected waves was modelled by relative time delays such as  $\Delta T$  (time between S and D)<sup>41</sup>, crest time (CT, time from onset to S)<sup>15</sup>,  $T_{Sys}$  (time in systolic phase), and  $T_{Dia}$  (time in diastolic phase), and  $T_{Ratio}$  ( $T_{Sys}/T_{Dia}$ )<sup>42</sup>. Additionally, the time delay between the systolic peak and dicotic notch is related to the



**Figure 1.** Overview of the fiducial points detected and some of the features extracted from the PPG waveform for one individual (Male, Age: 24, BMI: 25.1) during the four stages of the study protocol: (a) rest, (b) dose increase, (c) maximum infusion, and (d) washout. Examples of the following features are provided: Crest time (CT),  $\Delta T$ , reflection index (RI),  $width_{25}$ ,  $width_{50}$ , slope transit time (STT), A1, A2, Gaussian estimation of the transit time of the reflected wave (Gauss RTT) and augmentation index (Gauss AI),  $slope_{b-c}$  and  $slope_{b-d}$ . Acronyms: S - Systolic peak, N - Dicotic notch, D - Diastolic peak, a-e - waves of the APG.

Category	Feature notation and publication	Description or formula
PPG morphology	Notch amplitude ( $N_{amp}$ )	Amplitude of the dirotic notch (see Fig. 1)
	Reflective index (RI) <sup>38</sup>	Amplitude of the diastolic peak (see Fig. 1)
	$\Delta T$ <sup>41</sup>	Time from systolic peak to diastolic peak (see Fig. 1)
	Crest Time (CT) <sup>15</sup>	Time from onset to systolic peak (see Fig. 1)
	$T_{Sys}$ <sup>42</sup>	Time in systolic phase (see Fig. 1)
	$T_{Dia}$ <sup>42</sup>	Time in diastolic phase (see Fig. 1)
	$T_{Ratio}$ <sup>42</sup>	$T_{Sys}/T_{Dia}$
	Slope transit time (STT) <sup>50</sup>	Slope of straight line from onset to peak (see Fig. 1)
	Stress-Induced Vascular Response Index (sVRI) <sup>39</sup>	$\mu_{Dias}/\mu_{Sys}$
	A1 <sup>40</sup>	Area under PPG in systolic phase (see Fig. 1)
	A2 <sup>40</sup>	Area under PPG in diastolic phase (see Fig. 1)
	Inflection point area (IPA) <sup>40</sup>	A2/A1
	Width <sub>25</sub> <sup>41</sup>	Width of the PPG at 25 % of its amplitude (see Fig. 1)
	Width <sub>50</sub> <sup>41</sup>	Width of the PPG at 50 % of its amplitude (see Fig. 1)
	Pressure index (PI) <sup>43</sup>	$\frac{t(N)-t(S)}{t(N)-t(W)} \times h$
	Normalised harmonic area (NHA) <sup>40</sup>	$\frac{\sum_{n=2}^N FFT^2(f_n)}{\sum_{n=1}^N FFT^2(f_n)}$
	Inflection and harmonic area ratio (IHAR) <sup>40</sup>	(1-NHA)/ IPA
	Skewness <sup>23</sup>	-
Kurtosis <sup>23</sup>	-	
VPG morphology	$Sys_{\mu}$ <sup>18</sup>	Mean of VPG in the systolic phase
	$Sys_{\sigma}$ <sup>18</sup>	Variance of VPG in the systolic phase
	$Dia_{\mu}$ <sup>18</sup>	Mean of VPG in the diastolic phase
	$Dia_{\sigma}$ <sup>18</sup>	Variance of VPG in the diastolic phase
APG morphology	$\frac{b}{a}, \frac{c}{a}, \frac{d}{a} \& \frac{e}{a}$ <sup>46</sup>	
	Ageing index (AGI) <sup>46</sup>	$\frac{b-c-d-e}{a}$
	slope <sub>b-c</sub> <sup>42</sup>	Slope of a straight line between $b$ and $c$ , normalised by $a$ (see Fig. 1)
	slope <sub>b-d</sub> <sup>42</sup>	Slope of a straight line between $b$ and $d$ , normalised by $a$ (see Fig. 1)
	PPG AI <sup>47</sup>	PPG augmentation index $d_{-2}/b_{-2}$
Gaussian decomposition	$A_{g1-4}, \sigma_{g1-4}, \mu_{g1-4}$	Amplitude, variance and mean of the four decomposed Gaussians
	Gaussian augmentation index (Gauss AI) <sup>12</sup>	$\max(g_s) - A_{g3}$ (see Fig. 1)
	Gaussian reflection index (Gauss RI) <sup>12</sup>	$\sum(g_s) - \sum(g_d)$
	Gaussian reflected wave transit time (Gauss RTT) <sup>48</sup>	$\mu_{g3} - \mu_{g1}$ (see Fig. 1)
	Gaussian augmentation index <sub>R</sub> (Gauss AI <sub>R</sub> ) <sup>48</sup> †	$\frac{A_{g1}-A_{g2}}{A_{g1}}$
	Gaussian reflection index <sub>R</sub> (Gauss RI <sub>R</sub> ) <sup>48</sup> †	$A_{g3}/A_{g1}$
	Gaussian approximation of left ventricular ejection time (Gauss LVET) <sup>51</sup>	See reference for definition
	Gauss <sub>Sys/Dias</sub> <sup>*</sup>	$\sum(g_s)/\sum(g_d)$
	Gauss <sub>A4/A1</sub> <sup>*</sup>	$A_{g4}/A_{g1}$
Gauss <sub><math>\sigma</math>4/A1</sub> <sup>*</sup>	$\sigma_{g4}/A_{g1}$	
PCA	PPG PCA <sub>1-3</sub> <sup>52</sup>	First 3 principal components of PPG beat
	VPG PCA <sub>1-3</sub> <sup>52</sup>	First 3 principal components of VPG beat
	APG PCA <sub>1-3</sub> <sup>52</sup>	First 3 principal components of APG beat

**Table 2.** Summary of features extracted from the PPG. \*Indicates features that, to the authors’ knowledge, have not been previously implemented for BP estimation. † Note that there are two Gaussian indices referenced reflection index and augmentation index. We refer to the second set using the subscript R to reflect the authors: Rubins et al.  $\mu_{Sys}$  and  $\mu_{Dias}$  represents the mean of the PPG during systole and diastole respectively.  $h$  is the height of the participant.  $t$  is the normalised time since the pulse onset. FFT is the fast Fourier transform.  $g_i$  refers to the  $i^{th}$  Gaussian;  $g_s = g_1 + g_2$  representing the systolic wave;  $g_d = g_3 + g_4$  representing the diastolic wave.  $\sum(g_i)$  is the area under the  $i^{th}$  Gaussian.

transit time of the dirotic wave to the site of reflection and back. Shin et al.<sup>43</sup> suggest that this time delay may be calibrated to the individual by normalising by the time delay between the maximum derivative of the PPG and the dirotic notch, and multiplying the resultant by the individual’s height. This feature is labelled pressure

index (PI). Lin et al.<sup>44</sup> recently suggested that these time delay features may additionally reflect changes in CO. CO monitoring using the PPG has been demonstrated using the inflection and harmonic area ratio (IHAR)<sup>40</sup>.

A small number of features parameterise changes in the overall PPG morphology. For example, Width<sub>25</sub> and Width<sub>50</sub> track variations in the width of the PPG at 25% and 50% of the overall amplitude respectively<sup>41,45</sup>. Skewness and kurtosis both characterise the distribution of the PPG beat values<sup>23</sup>. Kurtosis is a measure of the relative “tailness” of the distribution. Skewness is a measure of the asymmetry of the data around the mean.

**VPG morphology.** We included four features derived from analysing the VPG morphology. In order to model the changing velocity of the pulse wave during the two main phases of the cardiac cycle, we included the mean and standard deviation of the VPG during systole (Sys<sub>μ</sub> and Sys<sub>σ</sub> respectively) and diastole (Dia<sub>μ</sub> and Dia<sub>σ</sub> respectively)<sup>18</sup>.

**APG morphology.** We included eight features derived from analysing the APG morphology. The amplitude of each APG wave, normalised by the amplitude of *a*, exhibits a moderate linear relationship to age<sup>46</sup>. *b/a* increase with age, whereas *c/a*, *d/a*, and *e/a* decrease with age. These relationships were implemented into a single feature using the ageing index (AGI)  $(b - c - d - e)/a$ <sup>46</sup>. In addition, Mok Ahn et al.<sup>42</sup> proposed the gradient of the straight lines joining the *b* wave to the *c* and *d* wave (slope<sub>b-c</sub> and slope<sub>b-d</sub> respectively) as additional indicators of age-related arterial stiffness. Finally, Pilt et al.<sup>47</sup> suggested that the PPG waveform index (PPG AI), computed as the ratio of PPG waveform coincident with the *d* wave of the PPG (labelled *d*<sub>-2</sub>) to the ratio of PPG waveform coincident with the *b* wave of the PPG (labelled *b*<sub>-2</sub>), reliably tracks changes in arterial stiffness.

**Gaussian decomposition.** Each PPG pulse (of unit amplitude and duration) was decomposed into the summation of four Gaussians<sup>48</sup>. This approach has the advantage of providing a representation of the PPG pulse without reliance on fiducial point detection. Additionally, it allows for the modelling of reflected wave interactions which are thought to have a Gaussian profile<sup>49</sup>.

For a PPG pulse  $\zeta$  we computed the modelled pulse,  $\zeta^{\text{Gauss}}$ , as:

$$\zeta^{\text{Gauss}}(t, \Theta) = \sum_{i=1}^4 g_i(t, \theta_i) \quad (1)$$

where  $g_i$  represents the  $i^{\text{th}}$  Gaussian component modelled as:

$$g_i(t, \theta_i) = A_{g_i} \times e^{-\frac{(t-\mu_{g_i})^2}{2\sigma_{g_i}^2}} \quad (2)$$

where  $t$  is the normalised time duration for the PPG pulse, and  $\theta_i$  is a vector,  $[A_{g_i}, \mu_{g_i}, \sigma_{g_i}]$ , containing the respective amplitude, mean and variance of each Gaussian.  $\Theta = [\theta_1, \theta_2, \theta_3, \theta_4]$  and thus Gaussian decomposition parameterises each PPG beat into 12 components. To determine the optimum value for  $\Theta$ ,  $\hat{\Theta}$ , we implemented a bounded Levenberg-Marquart optimisation algorithm to minimise the root mean squared error loss,  $L^{\text{Gauss}}$ , between  $\zeta$  and  $\zeta^{\text{Gauss}}$  given as:

$$L^{\text{Gauss}}(\Theta) = \sqrt{\frac{1}{N} \sum (\zeta - \zeta^{\text{Gauss}}(t, \Theta))^2}, \quad \hat{\Theta} = \underset{\Theta}{\text{argmin}}(L^{\text{Gauss}}(\Theta)) \quad (3)$$

The optimisation was bounded such that all parameters were positive and the amplitudes were all less than 1. Additionally,  $\{\mu_1, \mu_2, \mu_3, \mu_4\}$  must be a monotonically increasing sequence.  $L^{\text{Gauss}}$  is non-convex and therefore the optimised values were dependent on initial conditions. For the first beat, the initial conditions were:  $\theta_1 = [0.9, 0.2, 0.01]$ ,  $\theta_2 = [2/3, 0.4, 0.01]$ ,  $\theta_3 = [0.5, 0.6, 0.01]$ , and  $\theta_4 = [1/3, 0.8, 0.01]$ . These parameters resulted in an evenly distributed set of Gaussians with equal widths. The amplitudes were determined in order to match the contour of a typical PPG beat. To encourage continuity of parameters from beat-to-beat, we used the optimised parameters for the previous beat as initial seeds for the optimisation of the current beat. Following the work of<sup>53</sup>, we set the  $\text{SQL}_{\text{PPG}}$  of each beat to 0 if the value of  $L^{\text{Gauss}}(\hat{\Theta})$  for that beat was greater than 0.03.

From pulse decomposition analysis<sup>8</sup>,  $g_1$  represents the incident forward travelling pressure wave produced by left ventricular contraction.  $g_2$  represents the tidal wave<sup>9</sup>.  $g_2$  is typically observed during systole, therefore  $g_s = g_1 + g_2$  reflects the systolic component.  $g_3$  represents the dicrotic wave<sup>10</sup>.  $g_4$  represents the additional minor reflections and re-reflections in the systemic vascular structure with lower amplitude than the two main reflected waves<sup>54</sup>.  $g_3$  and  $g_4$  are both typically dominant in the diastole phase, therefore  $g_d = g_3 + g_4$  reflects the diastolic component.

Table 2 provides a summary of the Gaussian decomposition features used. Together with the values of  $\hat{\Theta}$ , we implemented various features derived from the Gaussian decomposition that have been previously proposed as indicators of arterial stiffness in the literature<sup>48,51</sup>. Additionally, through observations of Gaussian decomposition in our dataset, we propose three new features: the ratio of the systolic component to the diastolic component (Gauss<sub>Sys/Dias</sub>); the amplitude of the fourth Gaussian scaled by the amplitude of the first Gaussian (Gauss<sub>A4/A1</sub>); and the variance of the fourth Gaussian scaled by the amplitude of the first Gaussian (Gauss<sub>σ4/A1</sub>) (scaling by variance rather than the amplitude of the first Gaussian gave a less informative parameter). Figure 1 shows an example of Gaussian decomposition for a typical participant during the four main stages of the study protocol: rest, dose increase, maximum infusion, and washout. Figure 1 also presents examples of feature extraction for Gaussian estimation of the transit time of the reflected wave (Gauss RTT) and augmentation index (Gauss AI).

**Principal components.** Principal component analysis (PCA)<sup>55</sup> maps high-dimensional data to a lower dimension along orthogonal principal components. These principal components account for the majority of the variation in the original data and therefore highlight regions of significant change in the PPG, VPG and APG signals. We computed PCA features using the following steps:

1. Resample all good-quality beats (defined as an  $SQI_{PPG} > 0.8$ ) from the PPG, VPG and APG signals to be 100 samples in length using cubic spline interpolation.
2. Pool all resampled PPG, VPG and APG beats from all participants to form 3 matrices:  $\Psi_{PPG}$ ,  $\Psi_{VPG}$  and  $\Psi_{APG}$  respectively.
3. Mean normalise each  $\Psi$  matrix.
4. Perform PCA on each  $\Psi$  independently by computing the eigenvectors of the corresponding covariance matrix and extract the first 3 principal components that correspond to the largest eigenvalues.

We extracted 3 principal components as this was found empirically to explain more than 85% of the variation in the  $\Psi_{PPG}$ ,  $\Psi_{VPG}$  and  $\Psi_{APG}$  datasets. A visualisation of the computed PCA eigenvectors is provided in Supplementary Information figure SI: 1.

**Features from the ECG.** To preprocess the ECG waveform, we performed the following steps. To suppress the impact of baseline wander, the ECG was filtered using an 8<sup>th</sup>-order Butterworth IIR high-pass filter with a cut-off frequency of 0.5 Hz. To suppress power-line interference, a 2<sup>nd</sup>-order IIR notch filter with centre frequency at 50 Hz (the frequency of mains power in the UK) was used. We detected the QRS complex following the work of Pan and Tompkins<sup>56</sup> and assessed the quality of the ECG ( $SQI_{ECG}$ ) following the work of Li et al.<sup>57</sup>.

The features we extracted from the ECG are summarised in table 3. Features relating to complexity and entropy of the ECG have been previously proposed to track changes in BP<sup>28,58</sup>. These features quantify the level of regularity and unpredictability of fluctuations over a time series. Generally, a higher-level complexity indicates a more irregular dynamic system. A lower-level complexity indicates the presence of central trends or cyclical patterns. Changes in entropy of the ECG time series have been shown to track changes in heart rate variability (HRV) caused by myocardial ischaemia<sup>59</sup> and also denote periods of cardiac arrhythmia<sup>60</sup>. We provide full details of the algorithms used in Supplementary Information SI: 2.

**Pulse arrival time.** PAT has been shown previously to provide a beat-by-beat estimate of changes in arterial stiffness and therefore may be a good surrogate for BP<sup>1</sup>. PAT and its corresponding  $SQI_{PAT}$  was computed in the same manner as we have previously reported<sup>1</sup>. We used PAT in a baseline model to compare the performance of PPG and ECG features for BP estimation.

**Computing the reference BP values.** Measurements of SBP, MAP, and DBP using sphygmomanometer cuffs have known limitations depending on posture, cuff-inflation hypertension and cuff size<sup>62,63</sup>. The sphygmomanometer cuff used in our study was programmed to inflate every minute. However, errors in cuff inflation prevented the Philips monitor from registering an accurate estimate, resulting in a missed data-point in the recorded BP time series. Therefore, data from the cuff was both noisy and sparse. In order to reduce the impact of these sources of error, we processed the cuff data using a cubic smoothing splines<sup>64</sup> algorithm. This allowed for both filtering and interpolation of the noisy blood pressure readings to a new sampling frequency  $f_{BP}$  set as once per minute.

Let the  $i^{\text{th}}$  BP observation,  $y_i$ , at time  $t_i$  for a participant be modelled by the relation:

$$y_i = f(t_i) + \epsilon_i, \quad i = \{1, \dots, N_{\text{meas}}\} \quad (4)$$

where  $N_{\text{meas}}$  is the number of BP measurements recorded for the participant.  $\epsilon_i$  forms a sequence of independently distributed random variables with zero mean. Cubic smoothing splines define an estimate,  $\hat{f}$ , of  $f$  that equates to a cubic spline with knots (transition points) at  $\hat{f}(t_i)$ . At these transition point, the values of  $\hat{f}$ ,  $\hat{f}'$ , and  $\hat{f}''$  (where  $f'$  and  $f''$  denote the first and second derivative of  $f$  respectively) all match. The exact form of  $\hat{f}$  is determined by minimising a loss  $L^{\text{BP}}$ :

Feature notation and publication	Description or formula
Hjorth mobility <sup>28</sup>	Estimate of the signal's mean frequency (Equation SI 1)
Hjorth complexity <sup>28</sup>	Estimate of the signal's bandwidth (Equation SI 2)
Fractal dimension <sup>28</sup>	Computed using Higuchi's algorithm <sup>61</sup> with $k_{\text{max}} = 17$
Shannon entropy (SE) <sup>28</sup>	Uncertainty of information content based on probability distribution (Equation SI 3)
Approximate entropy (approxEnt) *	Quantifies regularities of signal
Sample entropy (sampEnt) *	approxEnt computed without self comparisons
Multi-level sample entropy (MSE) *	Course approximations of sample entropy, computed at scales 2, 4, 6, 8

**Table 3.** Summary of features extracted from the ECG. \* Indicates features that, to the authors' knowledge, have not been previously implemented for BP estimation. Full details of the algorithms and equations used for computing each feature can be found in Supplementary Information SI: 2.

$$L^{BP} = p \sum_{i=1}^{N_{\text{meas}}} (y_i - \hat{f}(t_i))^2 + \int_{t_1}^{t_{N_{\text{meas}}}} \hat{f}''(t)^2 dt \quad (5)$$

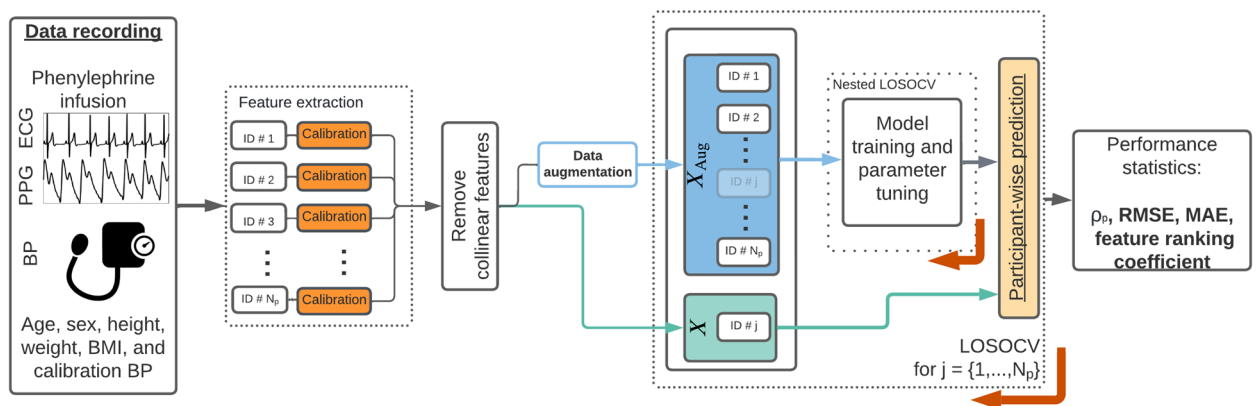
The first term of equation (5) imposes a penalty for the squared distance between observed BP values,  $y_i$ , and smoothed values  $\hat{f}(t_i)$ . The second term imposes a penalty for the complexity of  $\hat{f}$ , modelled as the accumulated second derivative of  $f$ .  $p$  is a constant that defines the relative weight placed on minimising the residual sum of squares against the complexity of  $\hat{f}$ . A very low value of  $p$  will result in the regressed function converging to a linear least squares estimate. A very high value of  $p$  will result in the smoothing spline converging to a cubic spline that passes through all data points.

As all participants were under the same protocol, we implemented a  $p$  value for SBP, MAP and DBP ( $p_{\text{SBP}}$ ,  $p_{\text{MAP}}$  and  $p_{\text{DBP}}$  respectively) that was common for all of them. Each respective  $p$  value was determined by extending the ordinary cross-validation strategy proposed in<sup>65</sup> by a grid search across the log-scaled range  $[10^{-3}, \dots, 10^8]$ . For each participant, the leave-one-out (LOO) RMSE was computed across the entire  $p$  range. The  $p$  value that minimised the participant-wise average LOO error was used.

**Estimating changes in BP.** We processed PPG features and PAT similarly to the methods proposed in<sup>1</sup>. This included: outlier detection to remove statistically significant deviations in values and a Kalman filter to reduce the effect of transient artefacts caused by noise. We then averaged the feature values within a window,  $w$ , of length 40s centred around each reference BP measurement (20s to the left, 20s to the right). Only beats of good quality, given by  $\text{SQI}_{\text{PPG}} > 0.8$  and  $\text{SQI}_{\text{PAT}} > 0.8$  respectively, were included in the window and if more than half of the window was deemed to be of bad quality, then the feature was not recorded for that window. We computed ECG features within the same window,  $w$ . If less than half of the window was deemed to be of good quality ( $\text{SQI}_{\text{ECG}} > 0.8$ ), then feature values were not recorded for that window. We handled missing data, caused by poor signal quality, by nearest neighbour imputation for each participant.

A schematic outlining our proposed steps for estimating  $\Delta\text{BP}$  is shown in Fig. 2. In this work, we adopted a *hybrid calibration*<sup>66</sup> approach in order to estimate  $\Delta\text{BP}$  using one of two ML regression models (LASSO+OLS or RF, defined in the sections below) as a function of an input feature set,  $X$ . We use  $x_i$  to represent the  $i^{\text{th}}$  observation of the feature set and  $x_{ij}$  to represent the datapoint of the  $i^{\text{th}}$  observation of the  $j^{\text{th}}$  feature. We implemented four different feature sets based on the different signals being analysed in this study. For each of the following groups we restricted  $X$  such that it includes features only from these sources:  $X \in \{\text{PPG}, \text{ECG}, \text{PPG}+\text{ECG}, \text{PAT}\}$ . We refer to these models and feature set combinations as LASSO + OLS<sub>PPG</sub> for a LASSO + OLS model with a PPG feature set, RF<sub>PPG + ECG</sub> for a RF model with PPG + ECG feature set, and so on.

We calibrated all feature and BP values to the individual participant using data recorded during the rest period of the study (first 5 minutes of the recording). We then removed all collinear features and implemented data augmentation to increase the size of the training set,  $X_{\text{Aug}}$ . We used nested leave-one-subject-out cross-validation (LOSOCV) to train, tune and evaluate the models. For each fold of the LOSOCV, one participant in turn was set as the test participant. Data from  $X_{\text{Aug}}$  set for all participants apart from the test participant were used to train and tune the models. Nested LOSOCV was used for model validation to optimise model hyperparameters ( $\lambda$  for OLS+LASSO and  $mtry$  for RF) with the aim of minimising RMSE. Data from  $X$  for the test participant was then used to evaluate the performance of the model. We used the following metrics to evaluate model performance:  $\rho_p$ , RMSE, and mean absolute error (MAE). Differences in performance metrics across all folds from the models



**Figure 2.** Schematic of the  $\Delta\text{BP}$  estimation pipeline for each of the proposed models. We extracted features from the PPG and ECG and averaged their values within a window of size 40s centred on times of cuff inflations. We then implemented a *hybrid calibration* approach such that the proposed models estimate  $\Delta\text{BP}$  from a baseline calibration value determined during the rest period. Data augmentation was implemented to increase the training and validation set size by interpolating between cuff inflations. Models were trained and evaluated in a nested leave-one-subject-out cross-validation (LOSOCV) framework shown here by the iterator  $j$  which indicates the test participant for that iteration. Participant  $j$  was then removed from the training/validation set ( $X_{\text{Aug}}$ ) for that iteration.



were evaluated for statistical significance by a two-tailed Wilcoxon signed-rank test. We adjusted the  $p$ -values for multiple comparisons using the Benjamini-Hochberg method<sup>67</sup>. This technique aims to control the number of type I errors (incorrectly rejecting the null hypothesis) by inflating the lowest  $p$  values (see<sup>67</sup> for more details).

**Calibration.** We adopted a *hybrid calibration*<sup>66</sup> approach to personalise the estimation models for each participant in our dataset. We use the superscript  $p$  to reflect an observation from the  $p^{\text{th}}$  participant. In this work, for each participant, we defined the baseline calibration value of BP ( $\widehat{\text{BP}}^p$ ) and the  $j^{\text{th}}$  waveform feature ( $\widehat{f}_j^p$ ) as their respective mean values during the 5-minute resting period at the start of the study. In the specific example of one participant for whom no BP readings were taken during the rest period, we used the mean values in a one-minute window centred on the first cuff inflation as the calibration readings.

For the  $p^{\text{th}}$  participant, BP values were subtracted from their calibration value to compute  $\Delta\text{BP}_i^p$ , as observations of the regression target vector. Additionally, the feature set  $X$  consisted of relative changes of each waveform feature from their calibration value<sup>68</sup>. Thus  $\Delta\text{BP}$  and  $x_{ij}^p$  took the form:

$$\Delta\text{BP}_i^p = \text{BP}_i^p - \widehat{\text{BP}}^p, \quad x_{ij}^p = \frac{f_{ij}^p - \widehat{f}_j^p}{\widehat{f}_j^p}, \quad i = \{1, \dots, N\}, \quad j = \{1, \dots, M\}, \quad p = \{1, \dots, N_p\}. \quad (6)$$

where  $N$  and  $M$  are the number of observations and features in  $X$  respectively.  $N_p$  is the number of participants.

*Hybrid calibration* strategies use participant demographics to personalise the model outputs. Thus, for all feature sets,  $X$ , we added participant age, sex, height, weight, BMI, and BP calibration ( $\widehat{\text{BP}}^p$ ) as static categorical features.

**Baseline reference.** For reference performance metrics, we implemented a simple baseline reference that assumed no BP changes for each participant from their baseline calibration value  $\widehat{\text{BP}}^p$  (i.e.  $\Delta\text{BP}_i^p = 0$  for all  $i$  and  $p$ ). We refer to this as *Baseline reference* and it indicates the minimum performance that must be achieved by the regression models.

**Removing collinear features.** Collinearity occurs when there is intercorrelation between multiple features<sup>69,70</sup>, thus violating the independent identically distributed (i.i.d.) assumption that is common in regression models. Additionally, the presence of collinearity inflates the variance of the regression parameters and makes it difficult to assess the importance of features. Collinearity in a feature set,  $X$ , can be highlighted by the condition number  $\kappa$  representing the ratio of the largest singular value of  $X$  to the smallest singular value. It can be computed as:

$$\kappa(X) = \|X^+\| \cdot \|X\| \quad (7)$$

where  $\|\cdot\|$  is the 2-norm of a matrix and  $X^+$  is the pseudo-inverse of the matrix  $X$ . Typically, a condition number greater than 30 is thought to indicate the presence of strong multi-collinearity in the dataset<sup>70</sup>.

It is likely that collinearity exists in the feature sets presented in this paper as there are multiple features describing similar characteristics, for example entropies of the ECG or time durations of the PPG. In order to remove the effect of collinearity and to allow for parsimonious models, we removed collinear features by investigating the variance inflation factor (VIF) defined as:

$$\text{VIF}_j = \frac{1}{1 - R_j^2} \quad (8)$$

where  $R_j^2$  is the unadjusted coefficient of determination for regressing the  $j^{\text{th}}$  feature on the remaining ones. If a feature can be accurately predicted using one, or more, other features then  $\text{VIF}_j$  would be large for that feature. Removing collinear features is an iterative process where on each iteration, the feature with the largest corresponding  $\text{VIF}_j$  is removed from the feature set until no features had a VIF greater than 10 (corresponding to  $R_j = 0.9$ )<sup>70</sup>.

**Data augmentation.** We implemented data augmentation in order to increase the feature set size by incorporating information from feature values between the reference BP cuff inflations. As shown in Fig. 2, we performed model training and validation on the augmented feature set, hereafter referred to as  $X_{\text{Aug}}$ , and performance metrics were reported using the original dataset  $X$ . We constructed  $X_{\text{Aug}}$  by interpolating between the reference BP values for each participant using the cubic smoothing splines outlined above at a new frequency,  $f_{\text{BP}} = 1/15$  Hz (four measurements per minute, as opposed to once a minute) with a smaller window size,  $w = 15s$ , to prevent overlapping windows violating the i.i.d. assumption.

**Regression models.** The models we implemented to estimate  $\Delta\text{BP}$  are outlined below. Different models were built for estimating  $\Delta\text{SBP}$ ,  $\Delta\text{MAP}$  and  $\Delta\text{DBP}$ .

**LASSO + OLS** To explore the linear relationship between each of the feature sets  $X$  and  $\Delta\text{BP}$ , ordinary least squares (OLS) linear regression was implemented. In order to prevent over-fitting and to improve model interpretability, we employed the Least Absolute Shrinkage and Selection Operator (LASSO) method to remove redundant features prior to linear regression. We refer to this model as LASSO+OLS. LASSO imposes the L1-norm

penalty to the residual sum of squares using non-negative values of shrinkage parameter  $\lambda$ . LASSO allows the removal of features by shrinking some feature coefficients,  $\beta$ , towards zero:

$$\beta = \operatorname{argmin}_{\beta} \left( \sum_{i=1}^N (Y_i - x_i \beta)^2 + \lambda \sum_{j=1}^M |\beta_j| \right) \quad (9)$$

where  $Y$  is the target vector ( $\Delta$ BP values) of length  $N$ , and  $M$  is the number of features. We optimised the  $\lambda$  hyperparameter by a nested LOSOCV loop. For each loop of the LOSOCV, LASSO feature selection was implemented and features with non-zero coefficients were used by OLS to compute  $\Delta$ BP estimates.

**Random forest** To explore potentially non-linear relationships between  $X$  and  $\Delta$ BP, we additionally built a Random Forest (RF) regression model. RF regression models utilise majority voting across multiple decision trees, each trained with a split criterion based on summed squared error (SSE)<sup>71</sup>. Each decision tree in an RF model was trained on a bootstrap of features. This approach reduces model variance whilst maintaining a low bias. As RF models select features upon training, we trained the model using all available features (i.e. without the need for LASSO). Typically, RF models are not very sensitive to choices in the number of trees ( $N_{\text{trees}}$ ), provided it is sufficiently high<sup>71</sup>. Therefore, we set the number of trees to 300. We optimised the number of features randomly selected for each node (labeled *mtry*) by a nested LOSOCV loop.

**SHAP values feature ranking coefficient.** A key objective of this work was to highlight features that have strong predictive power for estimating  $\Delta$ BP. We assessed the importance of each feature through Shapley additive explanation (SHAP) values<sup>72</sup>. Shapley values,  $\phi_j$ , represent the marginal contribution of each feature to individual model predictions. For a given regression model  $f(\cdot)$ , the marginal contribution of the  $j^{\text{th}}$  feature is computed through the difference in model outputs when trained with that feature compared to when trained without it,  $f(z_i \cup \{x_{ij}\}) - f(z_i)$ , where  $z_i$  is a subset of features in  $x_i$  not including  $x_{ij}$  ( $z_i \subseteq x_i \setminus \{x_{ij}\}$ ). Since the impact of adding a feature depends on the other features in the model,  $\phi_j$  is computed as the weighted sum of marginal contributions of the  $j^{\text{th}}$  feature over all possible subsets  $z_i \subseteq \{x_{11}, \dots, x_{iM}\} \setminus \{x_{ij}\}$ :

$$\phi_j(f, x_i) = \sum_{z_i \subseteq x_i \setminus \{x_{ij}\}} \frac{|z_i|!(M - |z_i| - 1)!}{M!} (f(z_i \cup \{x_{ij}\}) - f(z_i)) \quad (10)$$

where  $|z_i|$  is the number of features in  $z_i$ . As most models cannot handle missing data, SHAP values are computed by replacing  $f(\cdot)$  in equation (10) with  $f_x(\cdot)$ , a conditional expectation function of the original model. This step equates to replacing each missing feature with a random value representative of the datapoints the original model was trained over. Repeating this process, and averaging the results integrates out the missing value. It is demonstrated in<sup>72</sup> that SHAP values computed in this manner correspond to the only method of assessing feature importance that satisfies three desirable properties known as local accuracy, missingness, and consistency. A result of these properties is that for a complex, non-linear, model such as RF, a simpler explanation model  $g(\cdot)$  can be approximated through the linear combination of SHAP values across all features:

$$g(x_i) = \phi_0 + \sum_{j=1}^M \phi_j \quad (11)$$

where  $\phi_0$  is the expected model value over the training set. Equation (11) allows for local interpretability of model outputs by examining each feature's SHAP value for a given estimation. As SHAP values are computationally expensive to compute (there are  $2^M$  distinct coalitions of feature values), we computed approximates for the SHAP values through KernelSHAP (for LASSO+OLS) and TreeSHAP (for RF) implementations in the SHAP Python library<sup>72,73</sup>.

For each loop of the cross-validation (CV), we assessed the overall importance of each feature as the mean absolute SHAP value across the training data. To report the variability of the feature importance across the CV, we computed a ranking coefficient. The ranks of all features were determined at each fold and normalised by the total number of features (1 being the highest rank, 0 being the lowest). For each feature, the distributions of the ranking coefficients across all folds were analysed. This step allowed for each fold of the LOSOCV to contribute equally to assessing the overall importance, while also providing a fair comparison across SBP, MAP, and DBP.

## Results

**Clinical study.** Thirty volunteers were recruited for our clinical study. We discarded the data from four participants from the analysis. For three of these participants, the reference ECG waveform did not include any periods of high-quality data as a result of errors in the connection of the ECG electrodes. For one participant there were errors recording the BP cuff data. Therefore, 26 participants made up our dataset. The demographics of the participants in the study whose data was used for analysis are shown in table 4. All participants were healthy with a median BMI of 22.5 kg/m<sup>2</sup> and no history of cardiovascular disease. The median age of participants was 28 years and there was an even split of sexes (53.8% female). On average, we achieved an increase of 20 mmHg in SBP, with a maximum increase of 40 mmHg in a subset of participants.

**Removing collinear features.** Figure 3a shows the correlation matrix of the total feature set (PPG + ECG + demographics), including 77 features. There were a large number of features (58.5%) that were significantly correlated ( $|\rho_p| > 0.8, p < 0.05$ ) with at least one other feature. This indicates a high level of collinearity, high-

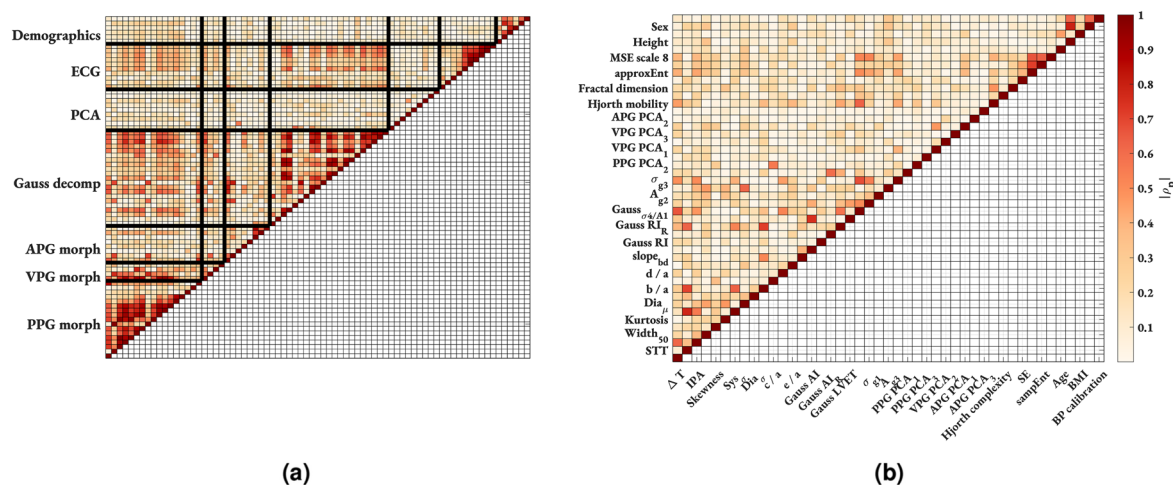
Descriptor	Value
Total number of participants	26
Male	12
Female	14
Average length of session (mins) <sup>1</sup>	28.0 (0.1)
Age (years) <sup>1</sup>	28.0 (9.0)
Height (cm) <sup>1</sup>	170.0 (18.0)
Weight (kg) <sup>1</sup>	69.5 (23.0)
Body Mass Index (kg/m <sup>2</sup> ) <sup>1</sup>	22.5 (5.2)
Reference maximum $\Delta$ BP per participant <sup>1</sup>	
$\Delta$ SBP (mmHg)	20.0 (8.0)
$\Delta$ MAP (mmHg)	17.0 (10.0)
$\Delta$ DBP (mmHg)	15.5 (9.0)
Reference $\Delta$ BP values <sup>2</sup>	
$\Delta$ SBP (mmHg)	6.5 (10.4)
$\Delta$ MAP (mmHg)	5.3 (8.8)
$\Delta$ DBP (mmHg)	4.7 (8.2)

**Table 4.** Demographics of the population in the clinical study. <sup>1</sup> Participant-wise median (IQR), <sup>2</sup> mean (standard deviation) across dataset.

lighted by a condition number  $\kappa$  of 315. Figure 3b shows the correlation matrix after removing all collinear features. The remaining dataset contained 45 features with a condition number  $\kappa$  of 11, suggesting independence of features and encouraging parsimonious models. For completeness, Supplementary Information table SI 3 provides a list of the remaining features and the subset of the total feature set with which they have a strong correlation. Additionally, Supplementary Information table SI 3 provides the correlation with  $\Delta$ SBP for each feature across the whole cohort and on a participant-wise basis.

**Comparing model performance.** Table 5 shows the median and interquartile range (IQR) performance statistics for all models computed over all 26 folds of the LOSOCV for  $\Delta$ SBP. The results for  $\Delta$ MAP and  $\Delta$ DBP are provided in the Supplementary Information table SI 1 and table SI 2 respectively. We note large RMSE and MAE values for the *baseline reference* indicating that participants experienced a significant change in their SBP values in response to the weight-based dosing of phenylephrine.

All models outperformed results obtained with the *baseline reference*, and all apart from LASSO + OLS<sub>ECG</sub> reported statistically significant  $p$  values indicating that consistent improvements in performance statistics were observed ( $p < 0.05$  for all). For the PPG, ECG and PPG+ECG feature sets, the RF model consistently achieved stronger performance metrics than LASSO + OLS. Statistically significant improvements were recorded only in RMSE and MAE ( $p < 0.05$  for all). The Wilcoxon signed rank test failed to reject the null hypothesis of equal



**Figure 3.** Results of removing collinear features in PPG + ECG feature set. Pairwise linear absolute Pearson's  $|\rho_p|$  correlation matrix. (a) pre and (b) post removal of collinear features. Blank spaces represent non-significant ( $p > 0.05$ ) or weak ( $|\rho_p| < 0.2$ ) correlations. For clarity of labelling on the post-collinear feature set (b), every other label is presented on the y-axis with the remaining labels on the x-axis.

Model Name	$\rho_p$	RMSE *	MAE *
Baseline reference	- (-)	10.46 (6.18)	7.40 (4.87)
LASSO+OLS <sub>PPG</sub>	0.85 (0.44)	6.19 (4.05)	5.41 (2.41)
RF <sub>PPG</sub>	<b>0.88 (0.26)</b>	5.75 (4.73)	<b>4.75 (3.66)</b>
LASSO+OLS <sub>ECG</sub>	0.64 (0.70)	7.48 (6.37)	6.54 (4.63)
RF <sub>ECG</sub>	0.69 (0.45)	6.79 (5.66)	5.28 (4.57)
LASSO+OLS <sub>PPG+ECG</sub>	0.82 (0.40)	7.99 (5.74)	6.30 (4.62)
RF <sub>PPG+ECG</sub>	0.86 (0.23)	<b>5.66 (4.76)</b>	4.86 (4.29)
LASSO+OLS <sub>PAT</sub>	0.83 (0.10)	6.09 (4.25)	4.91 (3.33)
RF <sub>PAT</sub>	0.80 (0.15)	6.39 (4.23)	5.21 (3.69)

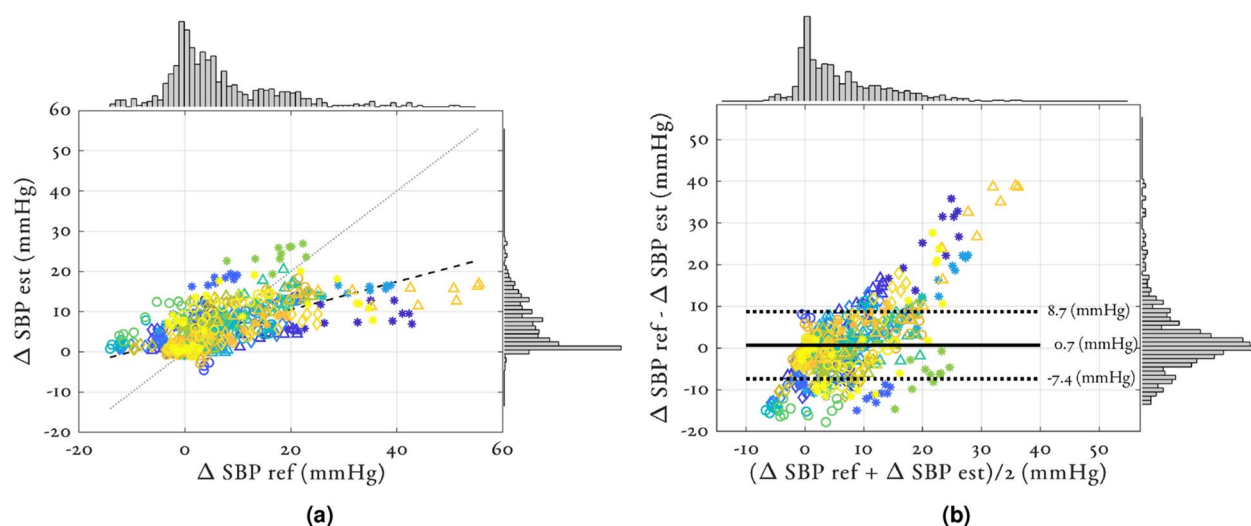
**Table 5.** Performance statistics of  $\Delta$ SBP estimation using the models proposed. Results are given as median (IQR) computed across all folds of the LOSOCV. Entries in bold indicate the best performance for that metric. \* results given in units of mmHg.

median  $\rho_p$  between LASSO+OLS and RF. The PPG feature set significantly outperformed the ECG feature set for all performance metrics and regression models. The absolute difference in the median  $\rho_p$ , RMSE, and MAE between RF<sub>PPG</sub> and RF<sub>ECG</sub> was 0.19 ( $p = 0.00007$ ), 1.04 mmHg ( $p = 0.005$ ) and 0.53 mmHg ( $p = 0.004$ ) respectively. RF<sub>PPG</sub> and RF<sub>PPG+ECG</sub> reported similar performance statistics with non-significant  $p$  values, indicating that adding ECG features to a feature set of PPG features offers little or no performance gain. For PAT, LASSO + OLS significantly outperformed RF across all performance metrics ( $p < 0.05$  for all). LASSO + OLS<sub>PAT</sub> achieved similar performance metrics to the RF<sub>PPG+ECG</sub>. LASSO + OLS<sub>PAT</sub> consistently resulted in the smallest IQR for all performance metrics.

Figure 4 shows the (a) correlation and (b) Bland-Altman plots for  $\Delta$ SBP estimation using the RF<sub>PPG+ECG</sub> model. Supplementary Information figures SI 2-3 show the correlation and Bland-Altman plots for  $\Delta$ MAP and  $\Delta$ DBP respectively. Individual participants are colour and marker-coded. The  $\rho_p$  value of the overall estimation was 0.64. The median participant-wise correlation coefficient was 0.86 with a range of 0.34 to 0.95. Figure 4a shows the histograms of the reference and estimated values. The reference  $\Delta$ SBP values ranged from -16.4 to 53.8 mmHg, but the estimated  $\Delta$ SBP values had a much tighter range of -3.37 to 22.2 mmHg. The bias of the overall error was 0.30 mmHg with a standard deviation of 8.05 mmHg (see Fig. 4b). We note also an additional bias where large values of  $\Delta$ SBP were underestimated and small values were overestimated. At peak infusion, the median (IQR) value of  $\Delta$ SBP across the cohort was 20 (8) mmHg (see table 4). We found large errors in the data for the four participants whose  $\Delta$ SBP at peak infusion exceeded 30 mmHg.

Supplementary Information figures SI 5-7 show the individual reference  $\Delta$ BP and estimated  $\Delta$ BP values using the RF<sub>PPG+ECG</sub> model across all participants in the study for SBP, MAP and DBP respectively.

**Feature importance.** To mitigate the variations in the training data between folds of the LOSOCV, overall SHAP value feature importances was assessed using a ranking coefficient. Figure 5 shows the median (across folds) RF and LASSO+OLS ranking coefficients for the PPG+ECG feature set for predicting SBP, with only the



**Figure 4.** Agreement between the reference  $\Delta$ SBP values from the sphygmomanometer cuff and the estimated  $\Delta$ SBP using RF<sub>PPG+ECG</sub>. Individual participants are colour and marker-coded. (a) The correlation analysis, the overall correlation was 0.64, the median participant-wise correlation coefficient was 0.86 with a range of 0.34 to 0.95. Black striped line shows regression line. Grey dotted line shows line of unity. (b) The Bland-Altman analysis, the bias of the overall error was 0.3 mmHg with a standard deviation of 8.05 mmHg.

top 15 features shown. In Supplementary Information figure SI 4, we show all features' median ranking coefficients for estimating SBP, MAP, and DBP. We quantified the agreement between the feature ranks for pairs of SBP, MAP, and DBP using the Kendall rank correlation coefficient,  $\rho_k$ <sup>74</sup>. It was found that the feature importance for SBP, MAP, and DBP estimation showed strong agreement with each other ( $\rho_k > 0.6$  for all) and as a result we continue to report SBP in the main body of this work. The top five features as determined by the median RF ranking coefficient were: kurtosis,  $\sigma_{g1}$ , Gauss $_{\sigma_4/A_1}$ , Hjorth mobility, and VPG PCA<sub>1</sub>. The top five features as determined by the median LASSO+OLS ranking coefficients were: kurtosis, Gauss $_{\sigma_4/A_1}$ , Gauss LVET, BP calibration, and IPA. Kurtosis had the highest feature importance in 13 out of 26 folds for both RF and LASSO + OLS (2<sup>nd</sup> highest in the remaining folds). In general, the ranking of features by LASSO+OLS followed the order of feature correlation values shown in Supplementary Information table SI 3. The only ECG feature that demonstrated significant importance in  $\Delta$ BP estimation was Hjorth mobility which had a median ranking coefficient of 0.93 for RF.

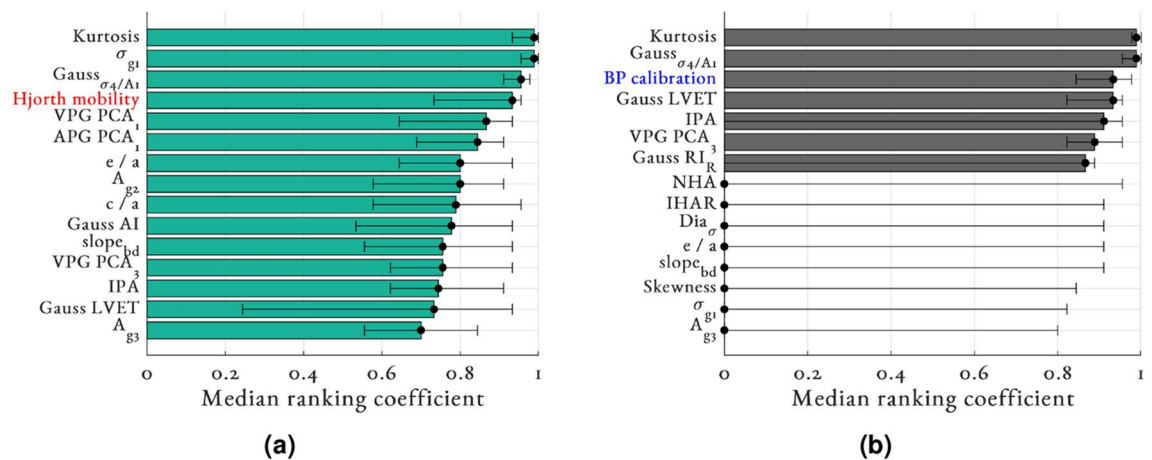
Figure 6 shows the relationship between  $\Delta$ SBP and the top 9 ranking, non-demographic, features from the RF<sub>PPG+ECG</sub> model. Different participants are coded by colour and marker to highlight clusters of features suggesting individual-specific feature changes. Both Pearson's  $\rho_p$  and Spearman's  $\rho_s$  correlation coefficient are provided. Spearman's correlation indicates monotonic (but not necessarily linear) relationships and so may provide further insight into the RF models (see for example the scatter of  $\Delta\sigma_{g1}$ ).

## Discussion

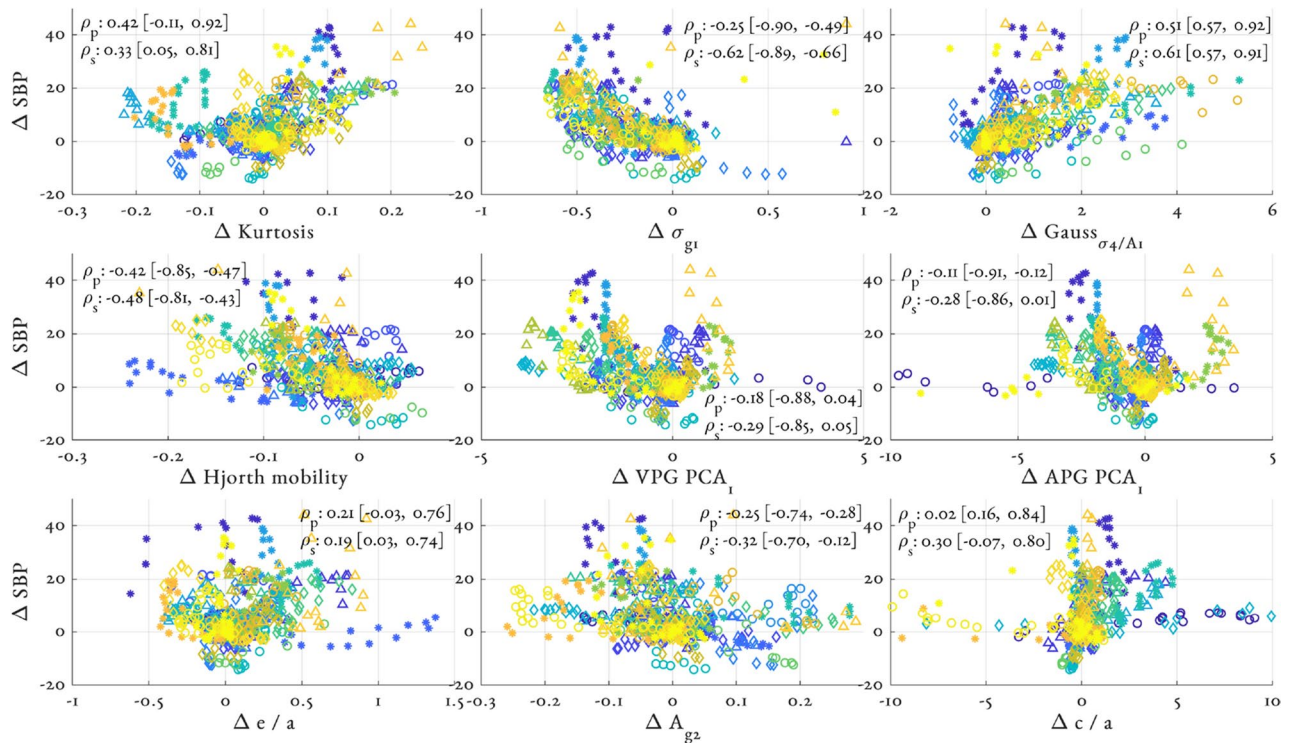
This work describes the methods for the non-invasive estimation of  $\Delta$ BP in healthy participants using morphological features from the PPG and ECG waveforms. Changes in BP were induced by the infusion of phenylephrine using a standard weight-based dosing protocol, instead of being BP-target driven (although BP was constantly under review by the clinician to ensure safety of the participants). One of the key advantages of this study was that, in a relatively short period of time, and while remaining supine and still, the participants experienced a wide range of BP values. This helped to validate algorithms for non-invasive, cuffless, estimation of SBP over a clinically useful range of  $\Delta$ SBP (-10 to 30 mmHg).

The PPG and the ECG waveforms offer great potential for non-invasive monitoring due to their ubiquity and ease of acquisition. The PPG in particular can be acquired using wearable devices such as smartwatches or using video plethysmography<sup>14</sup>. BP estimation using the PPG has been studied in a number of papers<sup>14</sup>, however the best features and models have remained unclear. A single-lead ECG may be recorded by three electrodes or via capacitive coupling<sup>75</sup>. The relationship between changes in BP and the ECG is governed by mechano-electric coupling (MEC). However, this connection is considered less robust to that of the relationship between the PPG waveform and BP. As a result, estimating  $\Delta$ BP from the ECG waveform has been explored in less detail in the literature.

**Observed changes in the PPG waveform morphology.** Figure 1 shows the changes in PPG beat seen for a typical participant during the four main stages of our study (a) rest, (b) dose increase, (c) max infusion, and (d) washout. This includes a loss of a clear dicrotic notch and a rising middle peak (often referred to as a tidal wave<sup>9</sup> and thought to be caused by reflected waves at the renal arteries<sup>8</sup>) that can overshadow the initial peak seen at rest. These changes have been reported previously<sup>18,41,49</sup> and are thought to be due to increasing amplitude and speed of reflected waves due to arterial stiffening. Additionally, they have also been shown to occur for age related arterial stiffening<sup>9</sup>. Note that in Fig. 1c at maximum infusion, the tidal wave has a peak greater than that of the original systolic peak (see Fig. 1b). As the systolic peak is almost universally defined as the maximum of the PPG beat<sup>37</sup>, this may lead to the tidal wave being incorrectly classified as the systolic peak. In which case,



**Figure 5.** Median SHAP values feature ranking coefficient for (a) random forest and (b) LASSO+OLS both with PPG + ECG feature set. Features are ordered by their respective median ranking coefficient. Error bars denote the range. Demographic features are highlighted in blue and ECG features in red for clarity. For brevity, only the top 15 features are provided here with the full list provided in Supplementary Information figure SI 4.



**Figure 6.** Relationship between top 9 ranking, non-demographic, features from the RF<sub>PPG+ECG</sub> model and  $\Delta$ SBP. Individual participants are colour and marker coded. Inter Pearson's  $\rho_p$  and Spearman's  $\rho_s$  correlation coefficient are shown as median and upper and lower quartiles. Additionally the 1st and 3rd quartiles of the participant-wise correlation coefficients are shown. Note that feature values have been both subtracted and divided by their value at baseline.

features such as CT,  $\Delta$ T, and STT, which rely on accurate systolic peak detection would be significantly affected. To account for this, we detected the systolic peak as the first turning point of the PPG pulse above its midpoint, as can be seen in Fig. 1c. We found that this definition was acceptable for our short, single perturbation study with minimal motion artefacts.

Other changes observed in the PPG waveform morphology during the phenylephrine infusion include variations in amplitude and time duration. Many factors can impact the amplitude of the PPG pulse, for example: respiration<sup>33</sup>; changes in the contact pressure of the pulse oximeter<sup>76</sup>; and changes in peripheral blood volume<sup>6</sup>. As these factors can occur independently of changes in blood pressure, and as recommended in<sup>76</sup>, we normalised the PPG pulse to have unit amplitude. Additionally, phenylephrine infusion activates the baroreflex response, increasing the parasympathetic nerve activity to the heart, and slowing heart rate down<sup>77</sup>. It was therefore decided to additionally normalise the time duration of the PPG pulse in order for features such as  $\Delta$ T (the time delay from systolic peak to diastolic notch) to be independent of HR.

**Model performance.** Typically, machine learning algorithms require a large amount of data in order to make robust estimations. In this work, we were limited by the number of participants in the study ( $N_p=26$ ) and by the number of data points per participant (typically once per minute for the 28 minutes of the recording per participant). In order to overcome these limitations, the models were trained, tuned and evaluated with a LOSOCV framework. Additionally, we implemented data augmentation to increase the number of data points in our training set by interpolating between the cuff measurements using cubic smoothing splines with an additional 3 augmented measurements every minute. Across all models, this data augmentation statistically significantly improved the performance statistics at the  $p < 0.001$  level, computed using a Wilcoxon signed rank test. Figure 3a splits all features into groups (PPG morphology, VPG morphology, APG morphology, Gaussian decomposition, ECG and demographics) and shows strong intra-group correlations indicating multi-collinearity in the dataset. We reduced feature collinearity by removing features with a VIF  $> 10$ . This increased the confidence of the model parameters, allowed for an appropriate examination of feature importance and encouraged parsimonious models.

For the PPG, ECG, and PPG+ECG feature sets, there were statistically significant improvements in RMSE and MAE when using the RF model compared to LASSO+OLS model at the  $p < 0.05$  level as computed by the corrected Wilcoxon signed rank test, whereas a non-significant difference was found in  $\rho_p$ . This indicates that LASSO+OLS was able to appropriately detect directional changes in  $\Delta$ SBP, although it was unable to determine the magnitudes of the changes. This points to a calibration issue whereby the LASSO+OLS model was unable to approximate the feature calibration gradients for each participant. RF models may have been better at estimating

the feature calibration gradients as these models allow for interactions between features and demographics. Indeed, larger overall importance was placed on demographic features by  $\text{RF}_{\text{PPG}+\text{ECG}}$  than  $\text{LASSO}+\text{OLS}_{\text{PPG}+\text{ECG}}$  (see Supplementary material figure SI 4). In linear models these static features can only influence intercepts and not gradients. Thus since the feature set was calibrated in order to predict  $\Delta\text{BP}$  from a baseline calibration value, low overall importance was placed on demographics by  $\text{LASSO}+\text{OLS}_{\text{PPG}+\text{ECG}}$ , with the exception of BP calibration. Supplementary material figures SI 5-7 demonstrates the marginal influence of demographics on the  $\text{RF}_{\text{PPG}+\text{ECG}}$  model estimates through the use of SHAP values. Demographics were found to have a varying influence on the  $\Delta\text{BP}$  estimation, with a positive impact on some participants (IDs: 003, 009 for example) and a negative impact on others (IDs: 018, 025 for example). Demographics are known to influence the PPG contour; in particular,  $c/a$  has been previously explored as a marker for age-related arterial stiffening<sup>41</sup>. Interestingly, the  $\text{RF}_{\text{PPG}+\text{ECG}}$  model puts little to no emphasis on the sex of the participant (see Supplementary Information figure SI 4) despite there being evidence of a sex-related dependency on the PPG morphology<sup>16</sup>. We would recommend in future studies that the interactions between participant demographics may be accounted for in a linear model by implementing a linear mixed-effects model with random effects for parameters such as age and BMI. We were unable to explore this line of work here due to the limited number of participants and their relative homogeneity.

Despite interactions with demographics, a significant calibration issue still persisted. Figure 4 shows the results of  $\Delta\text{SBP}$  estimation using the  $\text{RF}_{\text{PPG}+\text{ECG}}$  model via correlation and Bland-Altman analysis. Xing et al.<sup>78</sup> reported a Bland-Altman plot with similar characteristics where the model estimated  $\Delta\text{SBP}$  values to be in a much tighter range than was given by the reference and there were large errors found particularly at the high values. The largest errors found in all models occurred in the four individuals with the largest  $\Delta\text{SBP}$  at peak infusion (IDs: 002, 010, 023 and 026 in Supplementary Information figure SI 5). These four individuals experienced significantly larger values of  $\Delta\text{SBP}$  at peak infusion than the rest of the cohort ( $\Delta\text{SBP} > 33$  mmHg, whereas the median value across the cohort at peak infusion was 20 mmHg). The precise clinical effect of the weight-based dosing of phenylephrine in an individual would depend on the balance between their sensitivity to the increase in afterload, the effect of bradycardia on cardiac filling (therefore contractility, the Frank-Starling law<sup>79</sup>), and the proportion of venous/arterial action of phenylephrine causing the increase in preload and afterload. As a result, variations in  $\Delta\text{BP}$  at peak infusion were expected across the cohort. The median weight of the four outlier individuals (73 kg) was marginally larger than the cohort average (69.5 kg) and so on average a larger dose of phenylephrine would have been given. Other post-hoc assessments of demographics are not sufficient to distinguish these four individuals. Although in such a small cohort it is difficult to draw conclusions for these individuals, we suggest two possible explanations for the large errors found. Firstly, these individuals clearly experienced a significant change to their cardiovascular system in response to the dosing of phenylephrine and so the changes in the cardiovascular system may not be adequately represented by the features available. Secondly, the *hybrid calibration* strategy may be impacted by the small sample size. As a result, there will be difficulty in calibrating individuals that may be classed as outliers. In further work, to combat this inadequate calibration we suggest gathering a more longitudinal dataset for each individual and implementing *individual calibration* strategies.

Likely driven by the linear relationship between PAT and BP<sup>1</sup>,  $\text{LASSO} + \text{OLS}_{\text{PAT}}$  had stronger performance metrics than  $\text{RF}_{\text{PAT}}$ . PAT has been investigated as a surrogate measure of BP in this dataset previously<sup>1</sup> when we reported that *individual calibration*, as opposed to population-based *hybrid calibration* models, were needed for appropriate estimation of BP. In this work, there were insignificant differences found between  $\text{RF}_{\text{PPG}+\text{ECG}}$  and  $\text{LASSO} + \text{OLS}_{\text{PAT}}$  suggesting that features from the PPG and ECG have the same calibration constraint as PAT. Estimating BP from the ECG and/or PPG may hold two significant advantages over PAT. Firstly, the devices do not need to be perfectly synchronous and recorded on the same internal clock as is required for computing PAT estimates and has been reported as a limitation in certain datasets<sup>80</sup>. It would be possible to design a system for which BP was estimated from both devices when they are available and only from one device when the other was disconnected for any reason. Secondly, BP estimation from the PPG and ECG features was not impacted by the pre-ejection period (PEP) which is a significant limitation to BP estimation using PAT<sup>2</sup>. In our previous publication, we additionally reported on the influence of the pre-ejection period (PEP) on PAT estimates, noting that in this dataset PEP was found to have contributed between 28.8% and 35.2% of PAT<sup>1</sup>.

The results in table 5 suggest that PPG features have a significantly stronger relationship to changes in BP than the complexity features we extracted from the ECG. Additionally, despite Hjorth mobility demonstrating significant importance (median ranking coefficient = 0.93), when adding the ECG to a PPG feature set, no performance improvements were observed. Supplementary material figures SI 5-7 show the accumulated SHAP values for PPG features, ECG features, and demographics on the  $\text{RF}_{\text{PPG}+\text{ECG}}$  model estimates demonstrating that the model is largely dominated by the PPG component. As mentioned in the introduction, the theoretical relationship between changes in BP and the ECG is governed by MEC and the poor performance suggests that external, non-cardiac, control mechanisms may have a significant impact on the ECG-BP relationship, affecting the latter's ability to estimate  $\Delta\text{BP}$ . An additional explanation for the poor performance of the ECG feature set may be the choice of features used to explain the changes in ECG morphology; however to the authors' knowledge, there is no other work suggesting alternatives to ECG complexity features for BP estimation. For  $\text{RF}_{\text{ECG}}$ , improvements from a naive *baseline reference* assuming constant BP values were observed, but in general we suggest that ECG features on their own may not offer a viable solution to cuffless BP monitoring.

**Feature importance.** A key contribution of our work to the field of cuffless BP monitoring is a robust assessment of feature importance through the use of SHAP values and a ranking coefficient. We explored a large and comprehensive pool of features from both the PPG and the ECG gathered from a wide range of previous work (see table 3 and 2). Supplementary Information table SI 3 shows the features remaining for analysis after removing collinear features and the features from the original set with which they best correlate (defined

as  $|\rho_p| > 0.8, p < 0.05$ ). The overall correlations of these features to  $\Delta$ SBP across the cohort were in general quite low with only one feature (Gauss $_{\sigma_{4/A1}}$ ) having  $|\rho_{\Delta\text{SBP}}| > 0.5$ . There were a number of features that had significant participant-wise correlations (PWC) to  $\Delta$ SBP, 21 features with a median absolute PWC  $> 0.5$ . The disparity between low correlations across the cohort and high correlations on a participant-wise basis underpins the important need for *individual calibration* due to low intra-participant variability and high inter-participant variability.

BP is determined by CO and TPR, and changes in either of these may be represented by different features<sup>44</sup>. Phenylephrine causes a direct increase in TPR<sup>81</sup> via an increase in both arterial and venous vasoconstriction. Therefore, a large impact from the reflected waves, caused by impedance mismatches at points along the arterial tree (specifically the renal and iliac arteries<sup>8</sup>), was not unexpected. This was reflected in the observation that the majority of features that in this study have either a strong correlation or importance in estimating  $\Delta$ SBP, characterise the impact of the reflected pressure waves (Gauss $_{\sigma_{4/A1}}$ ,  $c/a$ , IPA). On the other hand, phenylephrine causes a mixed response in CO with the relationship governed by preload dependency<sup>81</sup>. The majority of participants in the study experienced a decrease in CO (see Supplementary Information figure SI 8), driven largely by a decrease in heart rate. Therefore, it is not surprising that at least one of the most important features ( $\sigma_{g1}$ ) represents changes in the upslope of the PPG which is driven by changes in CO<sup>44</sup>. However, it is not always clear how to link one feature to a specific BP control mechanism. Kurtosis (the feature with the highest importance for both RF and LASSO+OLS) for example, represents changes in the overall shape of the PPG.

We found some agreement between the features of the highest importance in our models to those reported in the literature. For example, features derived from the APG have been previously demonstrated to reflect age-related arterial stiffening<sup>42,46</sup>. In agreement with this, in our work,  $e/a$ ,  $c/a$ , and slope $_{bd}$  were found to have median ranking coefficients of 0.89, 0.84, and 0.75 respectively. In general, however, we suggest that there is little consensus on the appropriate features for BP estimation, both in our work and across the literature. There were large variations in the feature importances observed between folds (see Fig. 5). Additionally features such as STT, proposed by Addison et al.<sup>50</sup> and supported by a recent publication by Natarajan et al.<sup>68</sup>, had a very low median ranking coefficient of 0.16 in our dataset. Similarly, the best-performing features from Sun et al.<sup>18</sup> (Sys $_{\mu}$ , Sys $_{\sigma}$ , Dia $_{\mu}$ , and Dia $_{\sigma}$ ) and Miao et al.<sup>19</sup> ( $b/a$  and RI) all demonstrated poor performance in this work. Furthermore, to our knowledge, no previous works have suggested that Kurtosis may be a feature of significant importance for BP estimation.

A secondary explanation for the improved performance of the RF model relative to the LASSO+OLS model, may be that the features presented in this study (or at least the high-performing features presented in Fig. 5a) have a non-linear relationship to  $\Delta$ BP. This is further corroborated by the relationship between the top 9 ranking (non-demographic) features and  $\Delta$ SBP shown in Fig. 6. Only  $\sigma_{g1}$  shows a discernible global relationship to  $\Delta$ SBP. Individual, participant-specific clusters are apparent, highlighting the low intra-participant variability but high inter-participant variability. We additionally note that even within the clusters, non-linear relationships are often observed. This non-linear relationship is additionally supported by the results of Radha et al.<sup>12</sup> and Hasanzadeh et al.<sup>22</sup> who both reported performance improvements when estimating BP using a RF compared to a linear model. This significantly impacts the ability to develop *individual calibration* models using these features. For PAT, typically only 2 model parameters (a slope and an intercept) are required for accurate *individual calibration*. Therefore, theoretically, a dataset containing only two measurements of PAT and BP is required for accurate calibration to an individual (although in practice this number is much higher for accurate estimation of the model parameters, see<sup>1</sup>). Whereas, for non-linear modelling of the PPG or ECG features, many more parameters are required to be estimated thus forcing a much larger dataset requirement for accurate *individual calibration*.

It should be noted that many of the features used in this study were derived from fiducial points such as the dicrotic notch. Relying on fiducial point detection has a number of limitations for BP estimation. The detection algorithms often set arbitrary decisions or thresholds for fiducial point locations. As discussed previously, the typical definition for the systolic peak was not appropriate in this study due to the increasing influence of the tidal wave. Additionally, fiducial point detection algorithms will be valid up to a precision; small changes in BP observed in, for example, an ambulatory setting may result in very small perturbations in feature values that are indistinguishable from errors in fiducial point detection<sup>22</sup>. Finally, the fiducial points are not always detectable. The dicrotic notch has been reported to diminish in elderly individuals due to atherosclerosis (hardening of vessel walls and recruitment of collagen fibres to support walls)<sup>46,82</sup>. We found that the majority of the features of high importance (for example: kurtosis, PCA features and Gaussian decomposition features) did not require fiducial point detection. For the reasons stated above, these may offer more desirable representation of changes of the PPG.

**Limitations.** There are several limitations to the work presented in this study. Firstly, we perturbed BP via an infusion of phenylephrine, an  $\alpha_1$ -adrenergic receptor agonist that induces arterial and venous smooth muscle contraction<sup>30</sup>. In daily life, BP changes result from a diverse set of physiological mechanisms governed by the autonomic nervous system.  $\alpha_1$  receptors are typically activated in response to shock or low blood pressure<sup>83</sup>. Whereas, the resulting smooth muscle contraction is often activated during daily life activities, such as exercise<sup>84</sup>, in order to ensure adequate blood flow. Further work is required to understand how the relationship observed in this work compares to that present in daily life, however data acquired during daily life will be corrupted by motion artefacts.

Secondly, our results were reported across a small number of, relatively homogeneous, healthy participants ( $N_p=26$ ). We employed a *hybrid calibration* strategy to estimate changes in BP and utilised information from all available participants via a LOSOCV framework. However, this data-driven strategy requires a more heterogeneous cohort in order to improve model accuracy. This was particularly highlighted in the calibration issue for



the four individuals with the largest errors (see Supplementary Information figure SI 5). In particular, the cohort should contain participants across a wide range of age groups in order to account for variations in age-related arterial stiffness.

Thirdly, despite participants being administered a significant dose of phenylephrine (2mcg/kg/min), the changes we observed in the PPG (see Fig. 1) were very subtle. We were able to detect the fiducial points of the PPG accurately, as motion artefacts were reduced and the contact pressure of the pulse oximeter was maintained constant. However, in a real-world setting where such large variations in BP are uncommon and motion artefacts are a significant source of noise for PPG, this may be a significant limitation to BP estimation using PPG.

Finally, measurements of BP using a sphygmomanometer cuff are susceptible to various forms of noise that can distort the readings. The oscillometric device used as a BP reference in this study was compliant with the IEC 60601-2-30/EN60601-2-30 and with the American National Standard for Electronic or Automated Sphygmomanometers (ANSI/AAMI SP 10/92)<sup>85</sup> with a maximum mean error of  $\pm 5$  mmHg ( $\pm 0.7$  kPa) and a maximum standard deviation of 8 mmHg (1.1 kPa). The accuracy of the blood pressure cuff is a significant limitation to using single-point or *hybrid calibration* for BP estimation. Slight errors in a single cuff reading, caused by instrumentation error as well as user error (movement, wrong cuff size, etc) may translate into a consistent offset in BP estimation. Consider, for example, 006 in Supplementary Information figure SI 5, the initial calibration during the rest period sets, with both  $\Delta$ SBP cuff and  $\Delta$ SBP est at 0 mmHg. In the following 5 cuff inflations, the  $\Delta$ SBP cuff readings decreased to just under -5 mmHg, within the resolution of the ANSI/AAMI protocol. It is unclear whether this change in SBP is a real change (potentially caused by the participant relaxing after the start of the study) or if it was a result of instrumental errors in the blood pressure cuff. Either way, a consistent DC offset of 5–10 mmHg was observed for the remaining BP estimates in this individual.

## Conclusion

Under an infusion of phenylephrine, changes in the PPG (to a greater extent) and the ECG (to a lesser extent) reflect changes in BP that can be tracked using certain morphological features. For monitoring of BP by a single device, we recommend focusing on the PPG as this appears to be far superior to BP monitoring than using the ECG. In this study, we observed clear changes in the PPG in response to the dose increase of phenylephrine and characterised these by smooth muscle activation and a clear increase in the amplitude of the reflected tidal wave. These changes were mirrored in certain features and it appears that their relationship to  $\Delta$ BP may be non-linear. BP estimation using the PPG may offer similar performance to PAT which has significant limitations as it requires two synchronous devices (ECG and PPG) for accurate measurements. In general, the calibration protocol for accurate BP estimation requires more attention, especially if the relationship is non-linear. *Hybrid calibration* strategies may not adequately reflect the unique and individualised relationship between changes in BP and changes in the PPG. Therefore, they should be used with caution and only as a potential indicator of relative changes as opposed to a clinical assessment of BP.

## Data Availability

The datasets generated or analysed during the current study are not publicly available due to the sensitive and identifiable nature of our data, patient consent and restrictions of the ethics protocol to protect the privacy of patients involved in the study. Contact [eo.in.finnegan@eng.ox.ac.uk](mailto:eo.in.finnegan@eng.ox.ac.uk) for any queries.

Received: 11 October 2022; Accepted: 27 December 2022

Published online: 18 January 2023

## References

1. Finnegan, E. *et al.* Pulse arrival time as a surrogate of blood pressure. *Sci. Rep.* **11**, 1–21 (2021).
2. Payne, R. A., Symeonides, C. N., Webb, D. J. & Maxwell, S. R. Pulse transit time measured from the ECG: An unreliable marker of beat-to-beat blood pressure. *J. Appl. Physiol.* **100**, 136–141. <https://doi.org/10.1152/jappphysiol.00657.2005> (2006).
3. Mukkamala, R. *et al.* Toward ubiquitous blood pressure monitoring via pulse transit time: Theory and practice. *IEEE Trans. Biomed. Eng.* **62**, 1879–1901. <https://doi.org/10.1109/TBME.2015.2441951> (2015).
4. Sharma, M. *et al.* Cuff-less and continuous blood pressure monitoring: A methodological review. *Technologies* **5**, 21. <https://doi.org/10.3390/technologies5020021> (2017).
5. Peter, L., Noury, N. & Cerny, M. A review of methods for non-invasive and continuous blood pressure monitoring: Pulse transit time method is promising?. *Irbm* **35**, 271–282. <https://doi.org/10.1016/j.irbm.2014.07.002> (2014).
6. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Measur.* **28**, 1. <https://doi.org/10.1088/0967-3334/28/3/R01> (2007).
7. Mukkamala, R., Hahn, J.-O. & Chandrasekhar, A. Photoplethysmography in noninvasive blood pressure monitoring. In *Photoplethysmography*, chap. 11, 359–400 (Elsevier, 2022).
8. Baruch, M. C. *et al.* Pulse decomposition analysis of the digital arterial pulse during hemorrhage simulation. *Nonlinear Biomed. Phys.* **5**, 1–15 (2011).
9. Von Wowern, E., Östling, G., Nilsson, P. M. & Olofsson, P. Digital photoplethysmography for assessment of arterial stiffness: Repeatability and comparison with applanation tonometry. *PLoS ONE* **10**, e0135659. <https://doi.org/10.1371/journal.pone.0135659> (2015).
10. Nagasawa, T. *et al.* Blood Pressure Estimation by Photoplethysmogram Decomposition into Hyperbolic Secant Waves. *Appl. Sci. (Switzerland)* **12**, 1798. <https://doi.org/10.3390/app12041798> (2022).
11. Vybornova, A. *et al.* Blood pressure from the optical Aktiia Bracelet: A 1-month validation study using an extended ISO81060-2 protocol adapted for a cuffless wrist device. *Blood Press. Monit.* **30**, 0 (2021).
12. Radha, M. *et al.* Estimating blood pressure trends and the nocturnal dip from photoplethysmography. *Physiol. Meas.* **40**, 025006. <https://doi.org/10.1088/1361-6579/ab030e> (2019) (**1805.09121**).
13. Hosanee, M. *et al.* Cuffless Single-Site Photoplethysmography for Blood Pressure Monitoring. *J. Clin. Med.* **9**, 723. <https://doi.org/10.3390/jcm9030723> (2020).

14. Elgendi, M. *et al.* The use of photoplethysmography for assessing hypertension. *NPJ Digit. Med.* **2**, 60. <https://doi.org/10.1038/s41746-019-0136-7> (2019).
15. Millasseau, S. C., Kelly, R. P., Ritter, J. M. & Chowienczyk, P. J. Determination of age-related increases in large artery stiffness by digital pulse contour analysis. *Clin. Sci.* **103**, 371–377. <https://doi.org/10.1042/cs1030371> (2002).
16. Dehghanojamahalleh, S. & Kaya, M. Sex-related differences in photoplethysmography signals measured from finger and toe. *IEEE J. Transl. Eng. Health Med.* **7**, 1–7 (2019).
17. Boonya-Ananta, T. *et al.* Synthetic photoplethysmography (PPG) of the radial artery through parallelized Monte Carlo and its correlation to body mass index (BMI). *Sci. Rep.* **11**, 1–11 (2021).
18. Sun, S., Bezemer, R., Long, X., Muehlsteff, J. & Aarts, R. M. Systolic blood pressure estimation using PPG and ECG during physical exercise. *Physiol. Meas.* **37**, 2154–2169. <https://doi.org/10.1088/0967-3334/37/12/2154> (2016).
19. Miao, F. *et al.* A novel continuous blood pressure estimation approach based on data mining techniques. *IEEE J. Biomed. Health Inf.* **21**, 1730–1740. <https://doi.org/10.1109/JBHI.2017.2691715> (2017).
20. Wong, M. Y. M., Poon, C. C. Y. & Zhang, Y. T. An evaluation of the cuffless blood pressure estimation based on pulse transit time technique: A half year study on normotensive subjects. *Cardiovasc. Eng.* **9**, 32–38. <https://doi.org/10.1007/s10558-009-9070-7> (2009).
21. Mukkamala, R. & Hahn, J. O. Toward ubiquitous blood pressure monitoring via pulse transit time: Predictions on maximum calibration period and acceptable error limits. *IEEE Trans. Biomed. Eng.* **65**, 1410–1420. <https://doi.org/10.1109/TBME.2017.2756018> (2018).
22. Hasanzadeh, N., Ahmadi, M. M. & Mohammadzade, H. Blood pressure estimation using photoplethysmogram signal and its morphological features. *IEEE Sens. J.* **20**, 4300–4310 (2019).
23. Slapničar, G., Mlakar, N. & Luštrek, M. Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. *Sensors* **19**, 3420 (2019).
24. Schlesinger, O., Vigderhouse, N., Moshe, Y. & Eytan, D. Estimation and tracking of blood pressure using routinely acquired photoplethysmographic signals and deep neural networks. *Crit. Care Explor.* **2**, e0095. <https://doi.org/10.1097/ccce.000000000000095> (2020).
25. Reisner, A. T., Clifford, G. D. & Mark, R. G. The physiological basis of the electrocardiogram. *Adv. Methods Tools ECG Data Anal.* **1**, 25 (2006).
26. Timmermann, V. *et al.* An integrative appraisal of mechano-electric feedback mechanisms in the heart. *Prog. Biophys. Mol. Biol.* **130**, 404–417 (2017).
27. Reed, A., Kohl, P. & Peyronnet, R. Molecular candidates for cardiac stretch-activated ion channels. *Global Cardiol. Sci. Pract.* **2014**, 19 (2014).
28. Simjanoska, M., Gjoreski, M., Gams, M. & Bogdanova, A. M. Non-invasive blood pressure estimation from ECG using machine learning techniques. *Sensors (Switzerland)* **18**, 1160. <https://doi.org/10.3390/s18041160> (2018).
29. Harford, M. *et al.* Study protocol for an exploratory interventional study investigating the feasibility of video-based non-contact physiological monitoring in healthy volunteers by Mapping Of Lower Limb skIn pErfusion (MOLLIE). *BMJ Open* **10**, e036235 (2020).
30. Richards, E. & V. Maani, C. Phenylephrine. In *StatPearls [Internet]* (StatPearls Publishing, 2019).
31. NICE. Information for the public: High blood pressure (2011).
32. Villarroel, M. *et al.* Non-contact vital-sign monitoring of patients undergoing haemodialysis treatment. *Sci. Rep.* **10**, 1–21 (2020).
33. Charlton, P. H. *et al.* Breathing rate estimation from the electrocardiogram and photoplethysmogram: A review. *IEEE Rev. Biomed. Eng.* **11**, 2–20 (2017).
34. Chandrasekhar, A., Yavarimanesh, M., Natarajan, K., Hahn, J.-O. & Mukkamala, R. PPG sensor contact pressure should be taken into account for cuff-less blood pressure measurement. *IEEE Trans. Biomed. Eng.* **67**, 3134–3140 (2020).
35. Orfanidis, S. J. *Introduction to Signal Processing* (Prentice-Hall, Inc., 1995).
36. Elgendi, M., Liang, Y. & Ward, R. Toward generating more diagnostic features from photoplethysmogram waveforms. *Diseases* **6**, 20. <https://doi.org/10.3390/diseases6010020> (2018).
37. Charlton, P. H., Celka, P., Farukh, B., Chowienczyk, P. & Alastruey, J. Assessing mental stress from the photoplethysmogram: A numerical study. *Physiol. Measur.* **39**, 1. <https://doi.org/10.1088/1361-6579/aabe6a> (2018).
38. Padilla, J. M. *et al.* Assessment of relationships between blood pressure, pulse wave velocity and digital volume pulse. In *Computers in Cardiology*, vol. 33, 893–896 (IEEE, 2006).
39. Lyu, Y. *et al.* Measuring photoplethysmogram-based stress-induced vascular response index to assess cognitive load and stress. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 857–866 (2015).
40. Wang, L., Pickwell-MacPherson, E., Liang, Y. P. & Zhang, Y. T. Noninvasive cardiac output estimation using a novel photoplethysmogram index. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*, 1746–1749. <https://doi.org/10.1109/IEMBS.2009.5333091> (IEEE, 2009).
41. Elgendi, M. On the analysis of fingertip photoplethysmogram signals. *Curr. Cardiol. Rev.* **8**, 14–25. <https://doi.org/10.2174/157340312801215782> (2012).
42. Mok Ahn, J. New aging index using signal features of both photoplethysmograms and acceleration plethysmograms. *Healthcare Inf. Res.* **23**, 53–59. <https://doi.org/10.4258/hir.2017.23.1.53> (2017).
43. Shin, H. & Min, S. D. Feasibility study for the non-invasive blood pressure estimation based on ppg morphology: Normotensive subject study. *BioMed. Eng. Online* **16**, 1. <https://doi.org/10.1186/s12938-016-0302-y> (2017).
44. Lin, W.-H., Li, X., Li, Y., Li, G. & Chen, F. Investigating the physiological mechanisms of the photoplethysmogram features for blood pressure estimation. *Physiol. Meas.* **41**, 44003 (2020).
45. Awad, A. A. *et al.* The relationship between the photoplethysmographic waveform and systemic vascular resistance. *J. Clin. Monit. Comput.* **21**, 365–372. <https://doi.org/10.1007/s10877-007-9097-5> (2007).
46. Takazawa, K. *et al.* Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform. *Hypertension* **32**, 365–370. <https://doi.org/10.1161/01.HYP.32.2.365> (1998).
47. Pilt, K., Meigas, K., Ferenets, R., Temitski, K. & Viigimaa, M. Photoplethysmographic signal waveform index for detection of increased arterial stiffness. *Physiol. Meas.* **35**, 2027–2036. <https://doi.org/10.1088/0967-3334/35/10/2027> (2014).
48. Rubins, U., Grabovskis, A., Grube, J. & Kukulis, I. Photoplethysmography analysis of artery properties in patients with cardiovascular diseases. In *IFMBE Proceedings*, vol. 20 IFMBE, 319–322. <https://doi.org/10.1007/978-3-540-69367-3-85> (Springer, 2008).
49. Solà, J. & Delgado-Gonzalo, R. *The Handbook of Cuffless Blood Pressure Monitoring: A Practical Guide for Clinicians, Researchers, and Engineers* (Springer Nature, 2019).
50. Addison, P. S. Slope transit time (STT): A pulse transit time proxy requiring only a single signal fiducial point. *IEEE Trans. Biomed. Eng.* **63**, 2441–2444 (2016).
51. Couceiro, R. *et al.* Multi-Gaussian fitting for the assessment of left ventricular ejection time from the Photoplethysmogram. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 3951–3954. <https://doi.org/10.1109/EMBC.2012.6346831> (IEEE, 2012).
52. Xing, X. *et al.* Robust blood pressure estimation from finger photoplethysmography using age-dependent linear models. *Physiol. Meas.* **41**, 25007 (2020).

53. Wang, L., Xu, L., Feng, S., Meng, M. Q. & Wang, K. Multi-Gaussian fitting for pulse waveform using Weighted Least Squares and multi-criteria decision making method. *Comput. Biol. Med.* **43**, 1661–1672. <https://doi.org/10.1016/j.combiomed.2013.08.004> (2013).
54. Couceiro, R. *et al.* Assessment of cardiovascular function from multi-Gaussian fitting of a finger photoplethysmogram. *Physiol. Meas.* **36**, 1801 (2015).
55. Abdi, H. & Williams, L. J. Principal component analysis. *Wiley Interdiscip. Rev.: Computat. Stat.* **2**, 433–459. <https://doi.org/10.1002/wics.101> (2010).
56. Pan, J. & Tompkins, W. J. A Real-Time QRS Detection Algorithm. *IEEE Trans. Biomed. Eng.* **BME-32**, 230–236. <https://doi.org/10.1109/TBME.1985.325532> (1985).
57. Li, Q., Mark, R. G. & Clifford, G. D. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiol. Meas.* **29**, 15–32. <https://doi.org/10.1088/0967-3334/29/1/002> (2008).
58. Yang, S. *et al.* Blood pressure estimation with complexity features from electrocardiogram and photoplethysmogram signals. *Opt. Quant. Electron.* **52**, 135. <https://doi.org/10.1007/s11082-020-2260-7> (2020).
59. Leonarduzzi, R. F., Schlotthauer, G. & Torres, M. E. Wavelet leader based multifractal analysis of heart rate variability during myocardial ischaemia. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 110–113 (IEEE, 2010).
60. Li, H. *et al.* A new ECG signal classification based on WPD and ApEn feature extraction. *Circuits Syst. Signal Process.* **35**, 339–352 (2016).
61. Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Physica D* **31**, 277–283 (1988).
62. Ogedegbe, G., Agyemang, C. & Ravenell, J. E. Masked hypertension: Evidence of the need to treat. *Curr. Hypertens. Rep.* **12**, 349–355. <https://doi.org/10.1007/s11906-010-0140-4> (2010).
63. Lakhal, K., Ehrmann, S. & Boulain, T. Noninvasive BP monitoring in the critically ill: Time to abandon the arterial catheter?. *Chest* **153**, 1023–1039. <https://doi.org/10.1016/j.chest.2017.10.030> (2018).
64. Pollock, D. S. G. Smoothing with cubic splines (1993).
65. Craven, P. & Wahba, G. Smoothing noisy data with spline functions. *Numer. Math.* **31**, 377–403 (1978).
66. Mukkamala, R., Stergiou, G. S. & Avolio, A. P. Cuffless blood pressure measurement. *Ann. Rev. Biomed. Eng.* **24**, 1 (2022).
67. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* **57**, 289–300 (1995).
68. Natarajan, K. *et al.* Photoplethysmography fast upstroke time intervals can be useful features for cuff-less measurement of blood pressure changes in humans. *IEEE Trans. Biomed. Eng.* (2021).
69. Dormann, C. F. *et al.* Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).
70. Kim, J. H. Multicollinearity and misleading statistical results. *Kor. J. Anesthesiol.* **72**, 558 (2019).
71. Genuer, R., Poggi, J.-M. & Tuleau, C. Random Forests: some methodological insights. arXiv preprint [arXiv:0811.3619](https://arxiv.org/abs/0811.3619) (2008).
72. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30**, 1 (2017).
73. Lundberg, S. M. *et al.* From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2**, 56–67 (2020).
74. Kendall, M. G. A new measure of rank correlation. *Biometrika* **30**, 81–93 (1938).
75. Taji, B., Shirmohammadi, S., Groza, V. & Batkin, I. Impact of skin-electrode interface on electrocardiogram measurements using conductive textile electrodes. *IEEE Trans. Instrum. Meas.* **63**, 1412–1422 (2013).
76. Chandrasekhar, A., Yavarimanes, M., Natarajan, K., Hahn, J.-O. & Mukkamala, R. PPG sensor contact pressure should be taken into account for cuff-less blood pressure measurement. *IEEE Trans. Biomed. Eng.* **67**, 3134–3140 (2020).
77. Narkiewicz, K. *et al.* Baroreflex control of sympathetic nerve activity and heart rate in obstructive sleep apnea. *Hypertension* **32**, 1039–1043 (1998).
78. Xing, X. *et al.* An unobtrusive and calibration-free blood pressure estimation method using photoplethysmography and biometrics. *Sci. Rep.* **9**, 1–8 (2019).
79. Vincent, J. L. Understanding cardiac output. *Critical Care* **12**, 1. <https://doi.org/10.1186/cc6975> (2008).
80. Bennis, F. C. *et al.* Artifacts in pulse transit time measurements using standard patient monitoring equipment. *PLoS ONE* **14**, e0218784. <https://doi.org/10.1371/journal.pone.0218784> (2019).
81. Cannesson, M., Jian, Z., Chen, G., Vu, T. Q. & Hatib, F. Effects of phenylephrine on cardiac output and venous return depend on the position of the heart on the Frank-Starling relationship. *J. Appl. Physiol.* **113**, 281–289 (2012).
82. Millasseau, S. C., Ritter, J. M., Takazawa, K. & Chowienczyk, P. J. Contour analysis of the photoplethysmographic pulse measured at the finger. *J. Hypertens.* **24**, 1449–1456 (2006).
83. II, J. S. D. & Reddivari, A. K. R. Alpha 1 Receptor Agonists. *StatPearls* (2022).
84. Silfani, T. N., Chen, C., Patil, R. D. & DiCarlo, S. E. Vascular smooth muscle and exercise. *Sports Med. Train. Rehabil.* **8**, 301–320. <https://doi.org/10.1080/15438629909512535> (1998).
85. Stergiou, G. S. *et al.* A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement. *Hypertension* **71**, 368–374. <https://doi.org/10.1161/HYPERTENSIONAHA.117.10237> (2018).

## Acknowledgements

EF was supported by a EPSRC DTA Studentship. MV, SD, MH, JJ, PW and LT were funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

## Author contributions

Study design and conceptualisation was performed by M.H., P.W., L.T. and M.V. Data collection was performed by M.H., E.F., S.D., and M.V. E.F., M.V., S.D., and M.H. developed methodology and software. M.V., L.T., and P.W. provided supervision. E.F. prepared the first draft of this manuscript. All authors critiqued and edited the manuscript for intellectual content.

## Competing interests

LT and PW report significant grants from the National Institute of Health Research (NIHR), UK and the NIHR Biomedical Research Centre, Oxford, during the conduct of the study; modest grants and personal fees from Sensyne Health, outside the submitted work. EF, SD, MH, and MV declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-27170-2>.

**Correspondence** and requests for materials should be addressed to E.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023