Song *et al. Conn Health Telemed* 2024;3:300003
DOI: 10.20517/chatmed.2023.30

Connected Health And Telemedicine

Original Article

Open Access

Check for updates

An exploratory study of the relationship between pulse transit time and blood pressure based on causal inference

Xinyue Song, Xiaorong Ding

School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 611731, Sichuan, China.

Correspondence to: Prof. Xiaorong Ding, School of Life Science and Technology, University of Electronic Science and Technology of China, No.2006 Xiyuan Avenue Hi-Tech Zone (West), Chengdu 611731, China. E-mail: xiaorong.ding@uestc.edu.cn

How to cite this article: Song X, Ding X. An exploratory study of the relationship between pulse transit time and blood pressure based on causal inference. *Conn Health Telemed* 2024;3:300003. https://dx.doi.org/10.20517/chatmed.2023.30

Received: 11 Dec 2023 **First decision:** 10 Apr 2024 **Revised:** 24 Apr 2024 **Accepted:** 10 May 2024 **Published:** 15 May 2024

Academic Editor: Stefano Omboni **Copy Editor:** Yanbing Bai **Production Editor:** Yanbing Bai

Abstract

Aim: We propose to examine the causal relationship between the noninvasive features represented by pulse transit time (PTT) and blood pressure (BP), with the aim of mitigating the impact of confounding factor(s) and thus improving the performance of BP estimation.

Methods: We identified the causal graph of BP and the important noninvasive features extracted from electrocardiogram (ECG) and photoplethysmogram (PPG) via fast causal inference (FCI) algorithm, with the orientations of the edges in the causal graph being determined by the causal generative neural networks (CGNN) algorithm. With the knowledge obtained from the causal graph, we further used hierarchical regression model to estimate BP, and validated the proposed method on 17 subjects.

Results: We found that the obtained causal graph was almost consistent with the prior knowledge, and heart rate (HR) was one of the main confounding factors of PTT and BP. Incorporating HR into the hierarchical regression model to eliminate its confounding effect on the PTT-based BP estimation, the mean value of SBP and DBP estimation was improved by 1.27 and 1.89 mmHg, respectively, and the mean absolute difference (MAD) was improved by 2.28 and 3.60 mmHg, respectively.

Conclusion: Causal inference-based method has the potential to clarify the causal relationship between BP and related wearable noninvasive features, which can further shed light on developing new methods for cuffless BP with

acceptable accuracy.

Keywords: Cuffless blood pressure, causal inference, pulse transit time, heart rate

INTRODUCTION

Blood pressure (BP) is an important physiological parameter. Compared with traditional BP measurement techniques, the cuffless and continuous BP monitoring method based on wearable sensors has received extensive attention in recent decades because of its adaptability to long-term and high-comfort wear requirements. Since electrocardiogram (ECG) and photoplethysmogram (PPG) contain a large amount of cardiovascular information and are easily accessible from sensors, current algorithms for cuffless BP estimation mainly rely on these two signals. Some of these methods are based on the complex theory of arterial wave propagation, and construct empirical regression models for BP estimation through the noninvasive feature pulse transit time (PTT)[1,2]. Another class of methods extracts multiple noninvasive features from ECG and PPG, and then constructs data-driven models for BP estimation using deep learning and machine learning.

Whether it is an empirical regression model or a data-driven model, there are few practical applications for wearable BP monitoring devices[3]. The main reason for this is the inter-individual variability, which hampers the estimation accuracy, thus failing to meet the BP monitoring standards, such as the IEEE Standard for Wearable, Cuffless Blood Pressure Measuring Device[4] (IEEE 1708aTM-2019), and the European Society of Hypertension recommendations for the validation of cuffless blood pressure measuring devices[5].

PTT is considered the most promising feature in cuffless BP estimation, but previous studies found that the relationship between PTT and BP is highly individual- and condition-dependent[6-8]. The cardiovascular system is a complex dynamic system, and various hemodynamic parameters interact with each other. For example, changes in arterial stiffness will significantly affect the shape of PPG waveform, and also affect the degree of contraction and relaxation of blood vessels, thus affecting BP changes[9]. Therefore, there are likely to be confounding factors in the process of BP estimation, resulting in occasionally accurate and occasionally inaccurate estimation results.

In order to improve the accuracy of cuffless BP estimation, the causal relationship between BP and other relevant characteristics needs to be further clarified. Causality is different from correlation in its ability to better explain the dependence between variables and its greater reliability. Relying solely on correlation for data analysis may cause fallacies, as correlation does not imply causation[10]. The causal discovery algorithm can find the potential causal structure among variables and present it in the form of a causal graph, based on a given set of observational data[11].

From the perspective of causal inference, this paper attempts to clarify the causal relationship between BP and PTT, and explore whether there are confounding factors affecting BP estimation. Based on the obtained knowledge, we also investigate the possible way to develop an improved estimation model that could potentially address the confounding factors to provide new ideas for cuffless BP estimation.

METHODS

Dataset

The data for this study came from VitalDB, an open-access public dataset of patients' intraoperative vital signs and biosignals collected using the vital records program at the Department of Anesthesiology at Seoul National University Hospital[12]. It provides time-synchronized physiological data captured from various devices. We selected 120 samples that included original waveforms of arterial blood pressure (ABP), ECG, PPG, cardiac output (CO), and systemic vascular resistance (SVR).

Feature extraction

Figure 1A shows the details of feature extraction, which includes features commonly used in cuffless BP estimation such as HR, PTT, and PPG intensity ratio (PIR). PTT is defined as the time between the peak of the first derivative of PPG (dPPG)in the cardiac cycle and the peak of the ECG wave, and PIR has been proposed to be the ratio of minimum amplitude (I_L) to maximum amplitude (I_H) of a PPG signal in the cardiac cycle[13]. In addition, stiffness index (SI) and augmentation index (AI), which were proved to reflect the stiffness of blood vessels during pulse wave propagation, were extracted[14]. SI is defined as height divided by the time delay between systolic and diastolic peaks of the pulse wave. In order to locate the diastolic peak more accurately, we replace it with the f point of the second derivative of PPG (sdPPG). AI is the ratio of the peak systolic amplitude (x) to the peak diastolic amplitude of the pulse wave (y).

We also propose a new feature that indirectly reflects hemodynamics, called symmetry (SYM). Due to the presence of reflected waves, the PPG waveform of each cycle is not symmetrical, and the reflected wave speed is related to vascular aging, peripheral arterial compliance, arteriosclerosis, circumferential stress of the artery wall, *etc.*[15,16]. By analyzing the symmetry of PPG waveform, these factors can be reflected to a certain extent. The SYM is defined as:

$$SYM = \frac{DTW(Original\ PPG, Symmetrical\ PPG)}{N} \tag{1}$$

$$DTW(X,Y) = \min\left(\sqrt{\sum_{k=1}^{K} d_{w_k}}/K\right) \tag{2}$$

where dynamic time warping (DTW) is a method for calculating the similarity of two signals $X(1: m)$ and $Y(1: n)$. The distance $d(x_i, y_j)$ between each point of x_i of X and y_j of Y is computed and populated into a warping matrix D. The optimal warping path ($W = w_1, w_2, ..., w_k$) traverses from $D(1, 1)$ to $D(m, n)$, and d_{wk} is the distance at step k on the path W. As shown in Figure 1B, the *Original PPG* is the original PPG waveform of a single cardiac cycle, and the *Symmetrical PPG* is a symmetrical waveform obtained after the left half of the axis is flipped along the axis with the time of peak systole. The calculated DTW value is divided by the length of a single period PPG to get the normalized SYM.

The peak and trough values of ABP were extracted as reference systolic blood pressure (SBP) and diastolic blood pressure (DBP). Since the causal correlation between SBP and DBP is particularly strong, DBP is independent from other features when SBP is used as a conditional variable. Therefore, in the subsequent process of causal inference, mean arterial pressure (MAP) is used as the representative of BP, and MAP is calculated by adding two-thirds of DBP and one-third of SBP. Finally, the causal relationship between the other 8 features and MAP was analyzed.

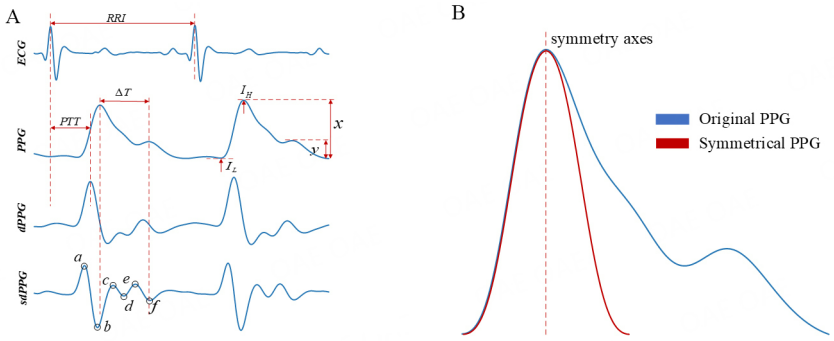

Figure 1. We show the details of feature extraction in Figure 1A and B. (A) parameters used in HR, PTT, PIR, SI, and AI calculations; (B) The extraction principle of feature SYM. HR: heart rate; PTT: pulse transit time; PIR: PPG intensity ratio; SI: stiffness index; AI: augmentation index; SYM: symmetry; PPG: photoplethysmogram; ECG: electrocardiogram; dPPG: derivative of PPG; sdPPG: second derivative of PPG; I_L: minimum amplitude; I_H: maximum amplitude.

Construct individual causality graph

Causal graphs are typically represented by directed acyclic graphs (DAG), where nodes represent random variables and directed edges represent causality from the tail to the arrow[17]. As shown in Figure 2, X is the cause of Y, and U is the common cause of X and Y. In a causal graph, if a variable is a direct cause of both a cause variable and an effect variable, then the variable is a confounding factor for the causal relationship[18].

To begin, use fast causal inference (FCI)[19] to obtain an initial causal graph for each individual. FCI algorithms perform a series of conditional independence tests to infer causal dependencies from time series data. The output result of the FCI algorithm is a Markov equivalence class, so there are nodes whose causal direction is not clear, and these nodes are represented by circles in the causal graph. Then, we use the causal generative neural network (CGNN)[20] algorithm to modify and locate the edges in the initial causal graph to obtain the final causal graph of the individual. CGNN utilizes the ability of neural networks to optimize DAG structures by minimizing the maximum mean difference between generated and observed data using a mountain climbing algorithm. For the final causal graph of each individual, we classify and encode the causal relationships to obtain the causal adjacency matrix of the individual:

$$A_{ij} = \begin{cases} 0 & x \quad y \\ 1 & x - - y \\ 2 & x \longrightarrow y \end{cases} \tag{3}$$

Where Aij represents the element in row i and column j of the adjacency matrix. In the final causality graph of an individual, if there is no line between x and y, it means that there is no direct causality between x and y, then $Aij = 0$; if there is an undirected edge between x and y, it indicates that there is a potential causal relationship between x and y, and it is necessary to add unmeasured potential variables to further clarify the direction of causality, making $Aij = 1$; if x is at the tail of an edge and points to y, then x is the direct cause of y, making $Aij = 2$.

Construct the population causal graph

Because there are individual differences, the causal graph obtained by different individuals will be different.

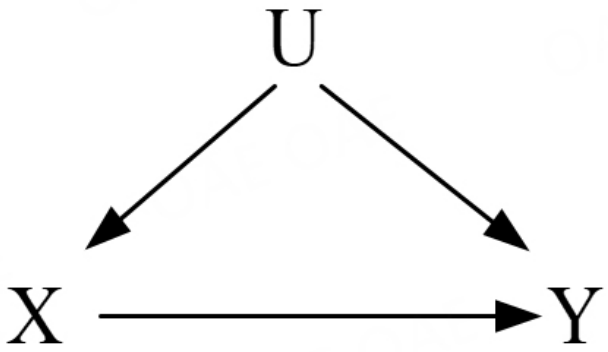

Figure 2. Example of the DAG. DAG: directed acyclic graphs.

In order to study the general causal relationship between BP and other relevant variables in the population, we need to retain the difference information among individuals as much as possible, so the causal adjacency matrix of 120 individuals is matrix accumulated to obtain the overall causal adjacency matrix of the population. According to the coding rules of the individual adjacency matrix, it can be considered that the value of each element in the population causal adjacency matrix can reflect the strength of the causal relationship between the variable *i* and the variable *j* to a certain extent. The above quartile is taken as the threshold, and MAP is taken as the outcome node of the causality map. The overall causality map of the population is drawn by selecting edges larger than the threshold.

Cuffless BP estimation model with knowledge obtained from the causal graph

In statistics, the method of eliminating the influence of confounding factors is usually to group the data according to the interval in which the confounding factors are located, and then re-study the relationship between variables within the group. By grouping PTT and BP according to the interval of the confounding factor *C*, we constructed a hierarchical PTT-based BP regression model with reference to the BP regression model of Proença *et al.*[21]:

$$SBP = \begin{cases} A_1 \cdot \ln PTT + B_1, & C \geq C_0 \\ A_2 \cdot \ln PTT + B_2, & C < C_0 \end{cases} \quad (4)$$

RESULTS

Population causal graph

Figure 3A shows the causal adjacency matrix of the resulting population sample. It can be seen that all the features we selected, except SI, have a causal value of more than 25 for MAP, indicating that they are the main causes of MAP and can be used as features for cuffless BP estimation. With MAP as the result node, edges greater than 25 are selected successively to draw the parent node of MAP and the parent node of the parent node according to the relationship in the matrix. If there is a bidirectional pointing edge, the one with the larger value is selected. The process is shown in Figure 3B, and finally, the overall causal graph is obtained in the fourth step. The following conclusions can be drawn from the overall causal graph:

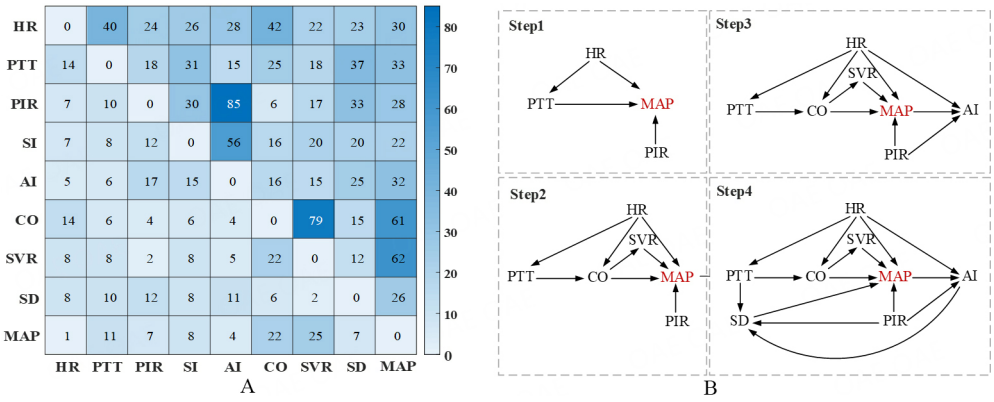

Figure 3. We present the obtained population causal adjacency matrix and the process of constructing the overall causal graph in Figure 3A and B. (A) the obtained population causal adjacency matrix. The value in the matrix represents the causal strength of the corresponding variable in the row to the corresponding variable in the column; (B) According to the threshold value, variables with fewer parent nodes are selected to be added to the causal graph in turn, and the overall causal graph is obtained in the fourth step.

(1) From the causal adjacency matrix, the causality values of CO and SVR are the largest among the direct causes of MAP, indicating that CO and SVR are the main causes of MAP. This is consistent with physiological knowledge, as MAP can be calculated as the product of CO and SVR[1,22]. When CO increases, the heart pumps more blood per minute, which leads to an increase in the flow of blood through the vessels, which increases MAP. When peripheral vascular resistance is increased, the flow of blood through the blood vessels is subjected to greater resistance, which causes MAP to rise because the heart requires greater force to overcome this resistance.

(2) Both SI and AI are used as indicators to reflect the stiffness of blood vessels, but the causal effect of AI and MAP is stronger for this group of subjects. In addition, our newly proposed feature SYM also shows causality with MAP, and SYM is the result of PTT, PIR, and AI at the same time, indicating that SYM can reflect the changes of these variables to a certain extent. The possibility of SYM for estimating BP can be further explored in the future work.

(3) It is evident in the diagram from step 1 of Figure 3B that HR is an important confounding factor in the estimation of BP using PTT. This has not been mentioned in previous studies, so we conducted a preliminary experiment to verify it.

Cuffless BP estimation

The data collected by Ding *et al.* were used for preliminary verification of the overall causal graph of the population obtained by the above method[23]. For 17 individuals, according to the HR change during deep breathing, the PTT, SBP, and DBP collected synchronously were divided into two groups according to the mean value of HR. Figure 4 shows the grouping of a subject. HR increases during inspiration due to the physiological reflex and neuroregulatory mechanism. PTT corresponding to high HR is divided into one group, and PTT corresponding to low HR is divided into one group for the subsequent input of the hierarchical model.

Figure 5 shows the BP estimates of a subject, where the green lines are the estimates of the original regression model and the yellow lines are the estimates of the hierarchical regression model. The gray lines represent the reference BP values, with the red and blue dots representing BP in the high HR group and the

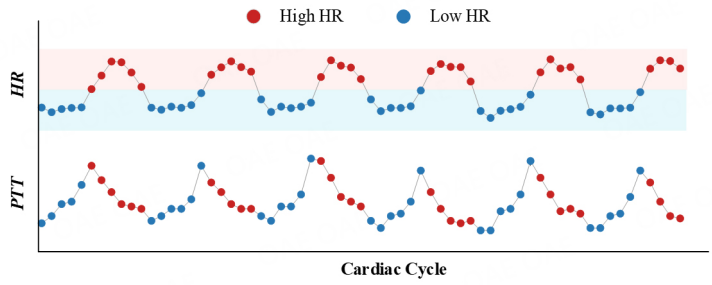

Figure 4. Take the data of a subject as an example, PTT were divided into two groups according to high heart rate and low heart rate. PTT: pulse transit time; HR: heart rate.

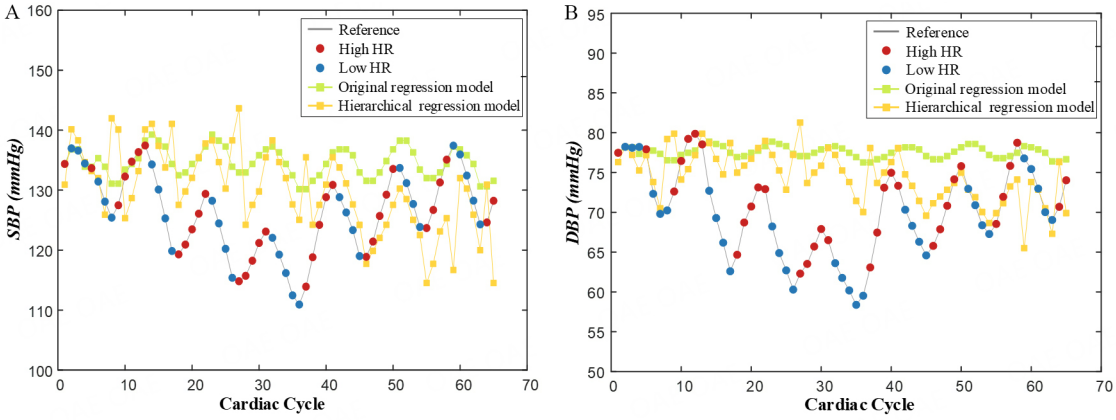

Figure 5. The blood pressure estimation results of the original regression model (green line) and the hierarchical regression model (yellow line) of SBP (A) and DBP (B) of a subject. SBP: systolic blood pressure; DBP: diastolic blood pressure; HR: heart rate.

low HR group, respectively. It can be seen that during 70 consecutive cardiac cycles, the reference BP fluctuated due to the modulating effect of respiratory activity. The change of BP value directly estimated by PTT without hierarchy is stable, and the trend of large changes in BP cannot be tracked, but this situation is optimized after hierarchy.

We conducted hierarchical regression for all 17 samples and averaged the results of all samples; the values of mean, standard deviation (SD); and mean absolute difference (MAD) for two models are shown in Table 1. Compared with the original regression model, the mean value of the hierarchical regression model in estimating SBP and DBP was improved to 0.49 and 0.84 mmHg, respectively, and the MAD was improved to 5.34 and 5.23 mmHg, respectively. These results, to some extent, confirm the conclusion of the causal graph that HR is a confounding factor for PTT estimation of BP, and the accuracy of BP estimation increases when PTT is grouped according to HR.

DISCUSSION

Causal inference can help explore the knowledge underlying the observed data. At present, the wearable or non-contact BP estimation models are based on correlation, ignoring the causality implied in the complex cardiovascular system, and it is difficult to obtain the characteristics that fundamentally cause BP changes, resulting in unsatisfactory BP estimation.

Table 1. Estimation accuracy of blood pressure by original regression model and hierarchical regression model

	SBP (mmHg)		DBP (mmHg)	
	Mean ± SD	MAD	Mean ± SD	MAD
Original regression model	1.76 ± 5.93	7.62	2.73 ± 6.75	8.83
Hierarchical regression model	0.49 ± 6.15	5.34	0.84 ± 5.44	5.23

SD: the standard deviation of the estimated error of all subjects; MAD: the mean absolute difference of the estimated error of all subjects; SBP: systolic blood pressure; DBP: diastolic blood pressure.

We found that HR is an important confounding factor in PTT estimation of BP. Although previous work has included HR as a parameter in multivariate regression models to predict BP based on PTT[6,24], they have not analyzed the influence of BP on the relationship between PTT and BP, so the accuracy is not ideal. In this work, HR is considered a confounding factor and a hierarchical PTT-based BP estimation model is established. The results obtained show that the BP estimation results are improved after stratification according to HR distribution. This is a revelation for most current methods of BP estimation, because the range of HR variation in different scenarios in people's daily life, such as work, sleep and exercise, can affect the effectiveness of wearable continuous BP estimation. This work has merely established a basic hierarchy of HR; how to avoid HR and other confounding factors in the PTT-based BP estimation model is the direction of our future research.

This study still has some limitations. First, there are not many hemodynamic parameters that can be directly measured, and the number of relevant samples that can be obtained is small, so some potential factors causing BP have not been analyzed. Secondly, due to the strong correlation between SBP and DBP, we chose a compromise approach, using MAP as a representative of BP values in causal inference. However, SBP or DBP was subsequently used as the output of the hierarchical model. In fact, MAP is somewhat different from SBP and DBP, and we will subsequently consider a more rational treatment of variables to solve the problem.

CONCLUSION

In this work, we extract representative characteristics of BP estimation in ECG and PPG, using causal discovery algorithms FCI and CGNN to explore the causal relationship between BP and its related characteristics in individuals, and then analyze the causal relationship of variables in the population on this basis. It is found in the causal graph that HR is a confounding factor in the estimation of BP by PTT. After layering the confounding factors HR, a hierarchical PTT-based BP estimation is constructed. Compared with the original regression model, the accuracy of BP estimation is improved by the hierarchical regression model. The other edges of the obtained cause graph indicate that there are many causes of BP, and these causes are also interrelated. How to use these causal relationships to improve the accuracy of BP estimation is worthy of further research.

DECLARATIONS

Authors' contributions

Made substantial contributions to the conception and design of the study and performed data analysis and interpretation: Song X, Ding X

Drafting and modification of manuscripts: Song X

Determine the final version of the manuscript: Ding X

Availability of data and materials
Not applicable.

Financial support and sponsorship
None.

Conflicts of interest
Both authors declared that there are no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

REFERENCES
1. Ding X, Zhang YT. Pulse transit time technique for cuffless unobtrusive blood pressure measurement: from theory to algorithm. *Biomed Eng Lett* 2019;9:37-52. DOI PubMed PMC
2. Barvik D, Cerny M, Penhaker M, Noury N. Noninvasive continuous blood pressure estimation from pulse transit time: a review of the calibration models. *IEEE Rev Biomed Eng* 2022;15:138-51. DOI PubMed
3. Chan G, Cooper R, Hosanee M, et al. Multi-site photoplethysmography technology for blood pressure assessment: challenges and recommendations. *J Clin Med* 2019;8:1827. DOI PubMed PMC
4. IEEE Standard Association. 1708-2014-IEEE standard for wearable cuffless blood pressure measuring devices. Available from: https://ieeexplore.ieee.org/document/6882122 [Last accessed on 14 May 2024].
5. Stergiou GS, Avolio AP, Palatini P, et al. European society of hypertension recommendations for the validation of cuffless blood pressure measuring devices: European society of hypertension working group on blood pressure monitoring and cardiovascular variability. *J Hypertens* 2023;41:2074-87. DOI
6. Kim JS, Kim KK, Baek HJ, Park KS. Effect of confounding factors on blood pressure estimation using pulse arrival time. *Physiol Meas* 2008;29:615-24. DOI PubMed
7. Chiang PH, Wong M, Dey S. Using wearables and machine learning to enable personalized lifestyle recommendations to improve blood pressure. *IEEE J Transl Eng Health Med* 2021;9:2700513. DOI PubMed PMC
8. Payne RA, Symeonides CN, Webb DJ, Maxwell SR. Pulse transit time measured from the ECG: an unreliable marker of beat-to-beat blood pressure. *J Appl Physiol* 2006;100:136-41. DOI PubMed
9. Finnegan E, Davidson S, Harford M, Watkinson P, Tarassenko L, Villarroel M. Features from the photoplethysmogram and the electrocardiogram for estimating changes in blood pressure. *Sci Rep* 2023;13:986. DOI PubMed PMC
10. Nogueira AR, Pugnana A, Ruggieri S, Pedreschi D, Gama J. Methods and tools for causal discovery and causal inference. *WIREs Data Min & Knowl* 2022;12:e1449. DOI
11. Hasan U, Hossain E, Gani MO. A survey on causal discovery methods for temporal and non-temporal data. Available from: https://arxiv.org/abs/2303.15027v2 [Last accessed on 14 May 2024].
12. Lee HC, Jung CW. Vital Recorder-a free research tool for automatic recording of high-resolution time-synchronised physiological data from multiple anaesthesia devices. *Sci Rep* 2018;8:1527. DOI PubMed PMC
13. Ding XR, Zhang YT, Liu J, Dai WX, Tsang HK. Continuous cuffless blood pressure estimation using pulse transit time and photoplethysmogram intensity ratio. *IEEE Trans Biomed Eng* 2016;63:964-72. DOI PubMed
14. Elgendi M. On the analysis of fingertip photoplethysmogram signals. *Curr Cardiol Rev* 2012;8:14-25. DOI PubMed PMC
15. Charlton PH, Paliakaitė B, Pilt K, et al. Assessing hemodynamics from the photoplethysmogram to gain insights into vascular age: a review from VascAgeNet. *Am J Physiol Heart Circ Physiol* 2022;322:H493-522. DOI PubMed PMC
16. Nichols WW. Clinical measurement of arterial stiffness obtained from noninvasive pressure waveforms. *Am J Hypertens* 2005;18:3S-10S. DOI PubMed
17. Zanga A, Ozkirimli E, Stella F. A survey on causal discovery: theory and practice. *Int J Approx Reason* 2022;151:101-29. DOI
18. Pearl J. Causal inference in statistics: an overview. *Statist Surv* 2009;3:96-146. DOI
19. Spirtes PL, Meek C, Richardson TS. Causal inference in the presence of latent variables and selection bias. Available from: https://arxiv.org/abs/1302.4983 [Last accessed on 14 May 2024].

20. Goudet O, Kalainathan D, Caillou P, Guyon I, Lopez-Paz D, Sebag M. Learning functional causal models with generative neural networks. Available from: https://arxiv.org/abs/1709.05321 [Last accessed on 14 May 2024].
21. Proença J, Muehlsteff J, Aubert X, Carvalho P. Is pulse transit time a good indicator of blood pressure changes during short physical exercise in a young population? *Annu Int Conf IEEE Eng Med Biol Soc* 2010;2010:598-601. DOI
22. Salvi P. Pulse waves: how vascular hemodynamics affects blood pressure. Springer Milan; 2012. p. i-138.
23. Ding X, Yan BP, Zhang YT, Liu J, Zhao N, Tsang HK. Pulse transit time based continuous cuffless blood pressure estimation: a new extension and a comprehensive evaluation. *Sci Rep* 2017;7:11554. DOI PubMed PMC
24. Cattivelli FS, Garudadri H. Noninvasive cuffless estimation of blood pressure from pulse arrival time and heart rate with adaptive calibration. In: 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks. IEEE; 2009:114-119.