



# Machine Learning based Predictors of Cardiovascular Disease among Young Adults

Dasharathraj K Shetty,<sup>1,#</sup> Lewlyn Lester Raj Rodrigues,<sup>1,#,\*</sup> Ajith Kumar Shetty<sup>2,#</sup> and Girish Nair<sup>3</sup>

## Abstract

The purpose of this research was to develop a data-driven model to test the association of physical, metabolic, and hemodynamic variables on the risk of cardiovascular disease. The structural equation modeling using the partial least square method has been adopted to analyze the data. A sample size of 685 young adults who were in sedentary, physically trained, and endurance-tested categories has been used in this research. Results have revealed that age and weight were the prominent predictors of cardiovascular disease among the physical variables, total glucose and triglycerides were the prominent predictors among the metabolic variables, and systemic vascular resistance and systolic blood pressure were the prominent predictors of cardiovascular disease among the hemodynamic variables. It was concluded that while all three variables are considered to be the antecedents of risk of cardiovascular disease, not all the parameters listed under these three categories have a statistically significant influence on the risk of cardiovascular disease. The results can be of use to medical practitioners as well as researchers in machine learning, as it adds to the repository of earlier studies and can be used by medical professionals in effective decision-making in disease prediction.

**Keywords:** Machine learning; Data-driven models; Cardiovascular disease; Physical variables; Metabolic variables; Hemodynamic variables; Structural equation modeling.

Received: 11 October 2021; Revised: 24 November 2021; Accepted: 26 November 2021.

Article type: Research article.

## 1. Introduction

Machine learning (ML) algorithms have been very promising in the application of artificial intelligence (AI) in medical sciences and are considered to be some of the essential diagnostic tools of the present day.<sup>[1]</sup> These algorithms have been very successful in diagnosing and predicting several diseases and have contributed to the enhancement of efficiency and effectiveness of the healthcare sector.<sup>[2,3]</sup> The field of ML has grown to such a level of maturity in a short span that medical education has also incorporated it as a field of study.<sup>[4,5]</sup> As a sub-field of AI, machine learning adopts several numerical techniques derived from optimization and statistics principles to enable the algorithm to learn

automatically by processing the data related to humans in a systematic and structured manner.<sup>[6-8]</sup>

Among the ML applications, statistics-based ML (also called statistical machine learning) is the specific technique adopted in this paper. Several research studies have used statistical machine learning, and the results seem to be very promising. Artificial neural network technique for a sample size of 683 was taken from the Wisconsin Diagnostic Data Set and predictive performance analysis was performed which showed very promising results.<sup>[9]</sup> The results indicated that the cell nuclei could be classified with an accuracy of up to 0.96, accompanied by a sensitivity of up to 0.99 and a specificity of up to 0.99. Song *et al.*<sup>[10]</sup> used a statistical ML technique for mortality prediction, and the model showed an accuracy of up to 0.89. While these are some studies that have used ML to predict the causes of diseases, research studies have found that statistical ML has not been very successful in predicting the antecedents of diseases.<sup>[11]</sup>

There are quite a good number of research studies in Cardiovascular disease (CVD) risk factors. This stream of research was initiated in the late 50s by the Framingham heart study (FHS).<sup>[12]</sup> Researchers have considered many different factors as antecedents to CVD, classified into background

<sup>1</sup> Department of Humanities and Management, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, 576104, India.

<sup>2</sup> Department of Anaesthesia & Critical Care, Sahyadri Narayana, Multispecialty Hospital, Shivamogga, Karnataka 576108, India.

<sup>3</sup> International Hospitality Management, Stenden University of Applied Sciences, Qatar.

<sup>#</sup> These authors contributed to this work equally.

\*E-mail: [l.rodrigues@manipal.edu](mailto:l.rodrigues@manipal.edu) (L. L. R. Rodrigues)

factors, metabolic factors, and hemodynamic factors.<sup>[13]</sup> There are several tools designed to measure these factors quantitatively in many different research studies. In the Framingham heart study based on a sample size of 3969 men and 4522 women, the Framingham tool<sup>[14,15]</sup> was used in which gender, age, diabetes mellitus (DM), smoking frequency (SF), systolic blood pressure (SBP), total cholesterol (TC), high-density lipoprotein (HDL) were considered to be the antecedents of CVD. The 12 pooled European studies were based on a sample size of 117,098 men and 88,080 women, and the scoring tool<sup>[16,17]</sup> was used, which considers age, gender, DM, SF, SBP, and TC. In the Scottish heart health extended Cohort prospective study based on a sample size of 6540 men and 6757 women, the assigned score tool<sup>[18]</sup> was used, which considers age, gender, DM, Rheumatoid arthritis (RA), family history of CVD (FHC), SF, SBP, TC, and HDL. A prospective study is based on a sample size of 18,460 men and 8515 women, and the PROCAM tool<sup>[19]</sup> that considers age, gender, SF, DM, body mass index (BMI), SBP, RA, TC, HDL, and FHC was used. In the QRESEARCH database with a sample size of 2.29 million, QRISK 2 tool<sup>[20]</sup> was used which considers age, gender, DM, Atrial fibrillation (AF), BMI, chronic kidney disease (CKD), SF, SBP, RA, FHC, TC, and HDL. In the modeling-based study of parameters in predicting the risk of CVD by the WHO, the WHO/ISH tool (WHO, 2007) was used, which considers age, gender, SMF, DM, SBP, and TC as the predictors of CVD. In the Women's Health study & Physician's health study II based on a sample size of 10,724 men and 24,558 women, the Reynolds tool used by Ridker considers age, gender, DM, SF, FHC, SBP, TC, HDL, and C-reactive protein (CRP). In the INTERHEART case-control study based on a sample size of 19470 individuals from many different countries, INTERHEART modifiable risk score tool<sup>[21]</sup> was used, which considers age, gender, DM, SF SBP, apoB/apoA1 ratio that represents the balance between atherogenic and anti-atherogenic lipoproteins, abdominal obesity, diet, physical and psychosocial factors. McGavock *et al.* have used a tool based on a sample size of 68 men and 67 women divided into three distinct groups of sedentary, physically trained, and endurance-trained individuals, the antecedents of CVD considered are age, gender, weight (kg), MBI ( $\text{kg}/\text{m}^2$ ),  $\text{VO}_2$  max ( $\text{mL}/\text{kg}/\text{min}$ ), cholesterol ( $\text{mg}/\text{dL}$  of blood), LDL ( $\text{mmol}/\text{L}$ ), HDL ( $\text{mmol}/\text{L}$ ), triglyceride (TGR) ( $\text{mmol}/\text{L}$ ), insulin ( $\mu\text{U}/\text{L}$ ), total glucose (GCS,  $\text{mmol}/\text{L}$ ), homeostasis model assessment (HMA), SBP ( $\text{mmHg}$ ), diastolic blood pressure (DBP) ( $\text{mmHg}$ ), heart rate (HR) ( $\text{bpm}$ ), systemic vascular resistance (SVR) ( $\text{dynes}\cdot\text{sec}\cdot\text{cm}^{-5}$ ) and flow-mediated dilation (FMD) (%). These are just some of the studies and the corresponding antecedents considered to be predictors of CVD.

Research studies have claimed that the causes of CVD are linked to the pre-birth period of the offspring, and several maternal conditions have been identified as the antecedents, which include smoking, sedentary lifestyle, obesity, and type of food intake.<sup>[22,23]</sup> Ho, Maddaloni & Buzzetti<sup>[24]</sup> performed a

literature review to study the relationship between obesity and CVD risk among the youth. The literature review indicated that the chances of developing a risk of CVD among the obese youth who have lost weight due to changes in lifestyle and the non-obese youth were the same. Some studies claim that overweight or obese have no association with CVD risk as long as they are metabolically healthy.<sup>[25-27]</sup> Mishra *et al.*<sup>[28]</sup> studied the association of gender, age, wealth, place of residence, educational level, marital status, hypertension, BMI, diabetes, and diet on CVD. The results revealed that females are at a higher risk of CVD, and it was observed that the disease starts in adolescence and aggravates in older age.

Through a meta-analysis of the literature, Steptoe & Kivimäki<sup>[29]</sup> found that work stress, social isolation, loneliness, and adverse metabolic changes influence CVD. Further, the stress was higher than the other variables on hypertension, which can lead to CVD.<sup>[30]</sup> An extensive literature review, which specifically focused on CVD, found that psychosocial stressors (mainly work stress) were among the major causes of CVD. Vasani *et al.*<sup>[31]</sup> used Framingham data of 3739 subjects and considered serum TC, SF, BMI, diabetes mellitus, SBP, DBP, and use of antihypertensive medicines as the antecedents of CVD in three different age groups, namely 60 years (women-729; men-522), 70 years (women-976; men-560) and 80 years (women-643; men-309) and found that among the antecedents considered, only BP was a consistent predictor of CVD irrespective of gender and age group. However, the findings also showed that a single measure of BP need not be a predictor of CVD, and it was inconclusive if time-averaged BP values could be a better predictor of CVD.

Age is considered an antecedent to CVD, as the heart's very functioning is subject to the life cycle operation.<sup>[32,33]</sup> In an extensive sample survey of 1,275, 174 adults (above 18 years), older people were more susceptible to contracting CVD.<sup>[33]</sup> Lloyd-Jones *et al.*<sup>[34]</sup> have found that in the absence of other traditional CVD risk factors, even elevated-age individuals had not shown the risk of CVD. So, the conclusion about the predictability of CVD risk with age as the independent variable appears to be inconclusive and more research is required in this direction.

People with increased BMI were likely to have CVD.<sup>[33]</sup> Maximal oxygen uptake ( $\text{VO}_2$  max, a measure of oxygen consumption during exercise) was associated with CVD. People with an uncontrolled SBP were more likely to have CVD than those with a controlled SBP.<sup>[33]</sup> In another study conducted in Eastern Finland with a sample size of 2682 men in the age group of 42-61 years Age, BMI, SBP, CLS, HDL, GCS, and LDL were associated with CVD.<sup>[35,36]</sup> Another study using a sample size of 320 belonging to two age groups of  $64 \pm 10$  and  $75 \pm 6$  with patient characteristics: hypertension, diabetes, dyslipidemia, beta-blockers, BMI ( $\text{kg}/\text{m}^2$ ), heart rate recovery, HRR ( $\text{bpm}$ ), SBP ( $\text{mmHg}$ ), DBP ( $\text{mmHg}$ ), pulse pressure ( $\text{mmHg}$ ), brachial-ankle pulse wave velocity (BAPWV) ( $\text{m}/\text{s}$ ), Creatinine ( $\text{mg}/\text{dL}$ ), LDL ( $\text{mg}/\text{dL}$ ), HDL ( $\text{mg}/\text{dL}$ ), triglyceride ( $\text{mg}/\text{dL}$ ) aimed to study the influence of

brachial-ankle pulse wave velocity on CVD. The study revealed that higher values of BAPWV were associated with CVD. Wolsk *et al.*<sup>[37]</sup> studied the influence of age, gender, and BMI on the hemodynamic parameters that were considered to be the antecedents to CVD with a sample size of 62 subjects (female-32; male-30). The results indicated that gender had no significant association with CVD risk, whereas BMI and BP had an association. It was also revealed that the hemodynamic parameters were better predictors of CVD risk.

Raffield *et al.*<sup>[38]</sup> investigated the association of age, sex, coronary artery calcified plaque (CAC), urine albumin creatinine ratio (UACR), insulin use, SF, and educational level with mortality due to CVD in type 2 diabetics using a sample size of 1022 European Americans. The results indicated that apart from age, CAC and UACR had the most significant association with CVD mortality. Further, the study results were validated using full sample cross-validation.

These research studies indicate that the risk of CVD is a multidimensional research construct. There are several antecedents and research is in progress throughout the world, considering many different combinations of the antecedents of CVD. Various tools have been developed to measure the antecedents of CVD, and it is observed that there were nine such widely used tools (Table S1). The McGavock tool had the highest number of predictors (15) used as antecedents for CVD, followed by the PROCAM tool (11), and QRISK 2 tool (10). On the other hand, three of the tools had the least number of predictors (6), namely the SCORE tool, WHO/ISH tool, and INTERHEART modifiable risk score tool. In terms of the frequency of use of the indicators, age, gender, and SBP were the most widely used indicators considered by all the tools, followed by Diabetes Mellitus and the smoking rate used by eight out of nine of the tools considered. Cholesterol and HDL Cholesterol followed the list with six and five of the tools using them, respectively. By and large, while there are commonly used antecedents, as mentioned before, the choice of an antecedent for the risk of CVD has been based on researchers' discretion, which could be based on the patient observations and experience of the medical professionals. However, it can be observed that age, gender, diabetes mellitus, smoking rate, family history of CVD, cholesterol, HDL Cholesterol (mmol/L), GCS, and Low-Density Lipoprotein (LDL) Cholesterol (mmol/L) are some of the common antecedents.<sup>[39]</sup>

A clear research gap that can be observed is the lack of information, on which one of the variables in three categories of variables of people has a bearing on the risk of CVD. Thus, this research specifically focuses on the influence of several antecedents above, categorized into physical variables, metabolic variables, and hemodynamic variables.

## 2. Materials and methods

### 2.1 The hypothetical model

A hypothetical model has been developed considering these studies, under three separate categories of human variables,

which act as the antecedents to CVD (Fig. S1).

Thus, the following hypotheses have been formulated to test the significance of the association between the antecedents and the risk of CVD.

H<sub>1</sub>: There is a significant association between age and the risk of CVD.

H<sub>2</sub>: There is a significant association between gender and the risk of CVD.

H<sub>3</sub>: There is a significant association between BMI and the risk of CVD.

H<sub>4</sub>: There is a significant association between VO<sub>2</sub> max and the risk of CVD.

H<sub>5</sub>: There is a significant association between weight and the risk of CVD.

H<sub>6</sub>: There is a significant association between cholesterol and the risk of CVD.

H<sub>7</sub>: There is a significant association between total glucose and the risk of CVD.

H<sub>8</sub>: There is a significant association between LDL cholesterol and the risk of CVD.

H<sub>9</sub>: There is a significant association between homeostasis mode assessment and the risk of CVD.

H<sub>10</sub>: There is a significant association between insulin and the risk of CVD.

H<sub>11</sub>: There is a significant association between LDL cholesterol and the risk of CVD.

H<sub>12</sub>: There is a significant association between TGR and the risk of CVD.

H<sub>13</sub>: There is a significant association between SVR and the risk of CVD.

H<sub>14</sub>: There is a significant association between SBP and the risk of CVD.

H<sub>15</sub>: There is a significant association between heart rate and the risk of CVD.

H<sub>16</sub>: There is a significant association between FMD and the risk of CVD.

H<sub>17</sub>: There is a significant association between DBP and the risk of CVD.

The literature review has enabled the identification of the research gap that has indicated that there is ample scope to identify the critical variables in each of the stated three categories of CVD predictors, and this has led to the postulation of the aforementioned hypotheses. The identification of these critical dimensions will be instrumental to the close monitoring of the parameters, to have better preventive control over the CVD. Framingham heart study (FHS)<sup>[12]</sup> based data has been simulated and used for statistical machine learning purposes leading to the testing of these hypotheses.

### 2.2 Study population

The data used in this research is from the patient data by McGavock, Anderson & Lewanczuk<sup>[40]</sup> who simulated a sample size of 685 based on the mean and standard deviation. The antecedents of CVD were to be studied for their

significance of influence on the risk of CVD under three distinct classes of physical, metabolic, and hemodynamic variables of the subjects. The range and mean of the parameters under consideration for the sample are indicated in Table S2.

**2.1. Mathematical background**

The basis for the hypothetical model is the Threshold Model proposed by Muthen. The dichotomous response to the presence or absence of CVD  $Y_{ijk}(k = 1,2...p)$  is a function of unobserved continuous response  $Y_{ijk}^*$  and can be mathematically given by Equation (1):

$$Y_{ijk} = \begin{cases} 1 & \text{if } Y_{ij}^* > \tau_k \\ 0 & \text{if } Y_{ij}^* \leq \tau_k \end{cases} \quad (1)$$

where  $\tau_k$  = is the threshold value.

Two distinct models operating at two levels are:

Level-1: The family model, based on the threshold model

At Level 1, the equations are in the linear form of the structural model as given in Equation (2).

$$Y_{ij} = \beta_i + KY_{ij}^* + \Gamma X_{ij} + \delta_{ij} \quad (2)$$

where  $\beta_i$  = a vector ( $p \times 1$ ) of  $i^{th}$  family;

$K = p \times p$  matrix of structural parameters;

$\Gamma X_{ij}$  = coefficient of relationship between latent response and predictor variables of  $i^{th}$  family,  $j^{th}$  individual;

$\delta_{ij}$  = is a vector of residuals ( $p \times 1$ ) of  $i^{th}$  family,  $j^{th}$  individual.

Level-2: Family random-intercept model, given in Equation (3).

$$\beta_i = \gamma + \rho_i \quad (3)$$

where  $\beta_i$  = a vector ( $p \times 1$ ) in the  $i^{th}$  family, and  $\gamma$  = a constant,  $\rho_i$  = a vector of residuals ( $p \times 1$ ) of  $i^{th}$  family.

The above Equations form Level-1 ( $j^{th}$  individual) and Level-2 ( $i^{th}$  family) for the structural equation modeling (SEM).

**2.2. Statistical machine learning**

The statistical machine learning (SML) methods commonly applied to remote sensing data are becoming popular these days for the simple reason that the use of statistical methods provides a proper direction in terms of utilizing, analyzing, and presenting the raw data available for Machine Learning.<sup>[40,41]</sup> In this research, the SML method has been used to test the hypothesis using Structural Equation Modelling as the tool.

**2.3. Structural equation modelling (SEM)**

The SEM has been the statistical tool used in this research. The tool's choice was mainly based on its proved predictive ability in many different contexts of research in medical sciences.<sup>[42,43]</sup> The method adopted in this research is to identify the CVD predictors. SEM can perform factor analysis and regression analysis simultaneously using the Partial Least Square Method (PLSM), which facilitates the addressing of multi-collinearity

effectively.

**3. Results**

**3.1 Measurement model**

The internal consistency of the data as measured through Cronbach's Alpha for physical, metabolic, and hemodynamic variables are in the range of 0.7-0.9, 0.6-0.9, and 0.7-0.9, respectively, which is in the acceptable range<sup>[44]</sup> (Tables S3, S4 & S5). The composite reliability estimates varying from 0.8-0.9 for all three variables indicate moderate to high- reliability values (Tables S3, S4 & S5). In this research, Rho-A, another measure of composite reliability uses values above 0.6 that are considered acceptable<sup>[45]</sup> (Tables S3, S4 & S5). Further, factor loading (FL) has been considered to measure the convergent validity of item reliability (IR). It is estimated as standardized values through factor reduction (Fig. 1 and Tables S6, S7 & S8) ranged from 0.7 to 0.9 for all the three variables above, which indicates a good correlation between the factor and the observed dimension. Another important test for the data's reliability is to ensure that each of the antecedents is mutually exclusive to ensure their distinctiveness. Discriminant Reliability is a widely accepted test for finding mutual exclusiveness by comparing square roots of the average variance extracted (AVE) for each dimension with the inter-item correlation of that dimension with the remaining dimensions. It is found that the AVE values are above the inter-item correlations, and hence, the variables chosen for the study are adequately mutually exclusive of each other (Tables S9, S10, and S11).

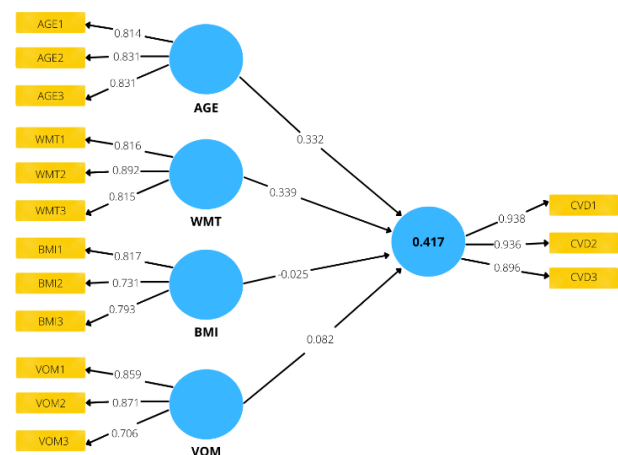


Fig. 1 Path model - physical variables.

**3.2 The structural model**

The R-squared values which indicate the percentage variance of dependent variables based on the variance in the independent variable for the physical, metabolic, and hemodynamic variables are 0.417, 0.4, and 0.529, respectively. These indicate that the model fit is acceptable (cutoff value 0.1)<sup>[46]</sup> (Table S4). The alpha values (p-values  $\leq 0.05$ ) (Tables 1, 2 & 3 and Figs. 2 to 6) indicate that the following hypotheses are supported:

**Physical variables**

H<sub>1</sub>: There is a significant association between age and the risk of CVD

H<sub>5</sub>: There is a significant association between weight and the risk of CVD.

**Metabolic variables**

H<sub>7</sub>: There is a significant association between total glucose and the risk of CVD

H<sub>12</sub>: There is a significant association between TGR and the risk of CVD

**Hemodynamic variables**

H<sub>13</sub>: There is a significant association between SVR and the risk of CVD.

H<sub>14</sub>: There is a significant association between SBP and the risk of CVD.

**Table 1.** t-test values - physical variables.

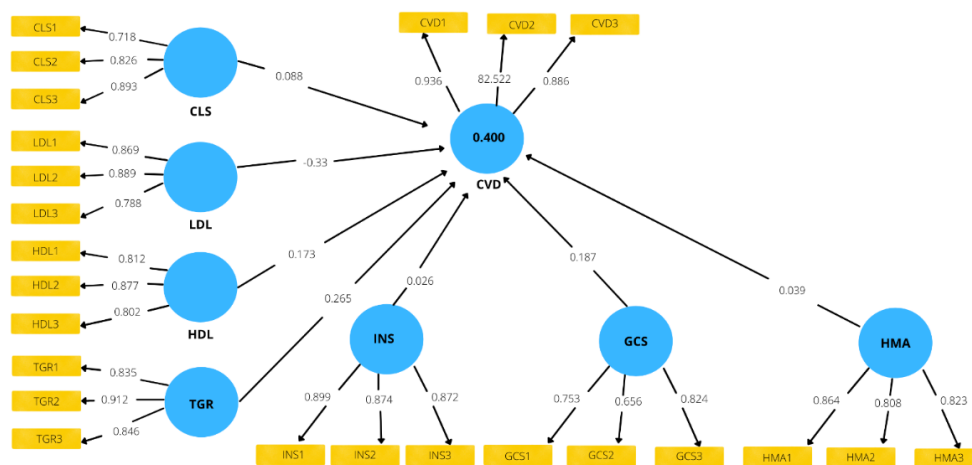
	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics ( O/STDEV )	P Values	Hypothesis
AGE -> CVD	0.33	0.33	0.08	4.21	0.00	Supported
BMI -> CVD	0.06	0.01	0.09	0.27	0.79	Not supported
VOM -> CVD	0.08	0.08	0.08	1.05	0.30	Not supported
WHT -> CVD	0.34	0.34	0.07	4.57	0.00	Supported

**Table 2.** t-test values - metabolic variables.

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics ( O/STDEV )	P Values	Hypothesis
CLS -> CVD	0.09	0.10	0.10	0.93	0.35	Not supported
GCS -> CVD	0.19	0.19	0.09	2.16	0.03	Supported
HDL -> CVD	0.17	0.17	0.11	1.55	0.12	Not supported
HMA -> CVD	0.04	0.04	0.08	0.47	0.64	Not supported
INS -> CVD	0.03	0.04	0.10	0.28	0.78	Not supported
LDL -> CVD	-0.03	-0.04	0.10	0.33	0.74	Not supported
TGR -> CVD	0.27	0.26	0.12	2.14	0.03	Supported

**Table 3.** t-test values - hemodynamic variables.

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics ( O/STDEV )	P Values	Hypothesis
DBP -> CVD	0.10	0.11	0.11	0.87	0.39	Not supported
FMD -> CVD	0.09	0.07	0.12	0.72	0.47	Not supported
HRR -> CVD	0.02	0.02	0.10	0.23	0.82	Not supported
SBP -> CVD	0.24	0.24	0.09	2.84	0.01	Supported
SVR -> CVD	0.38	0.40	0.12	3.29	0.00	Supported



**Fig. 2** Path model - metabolic variables.

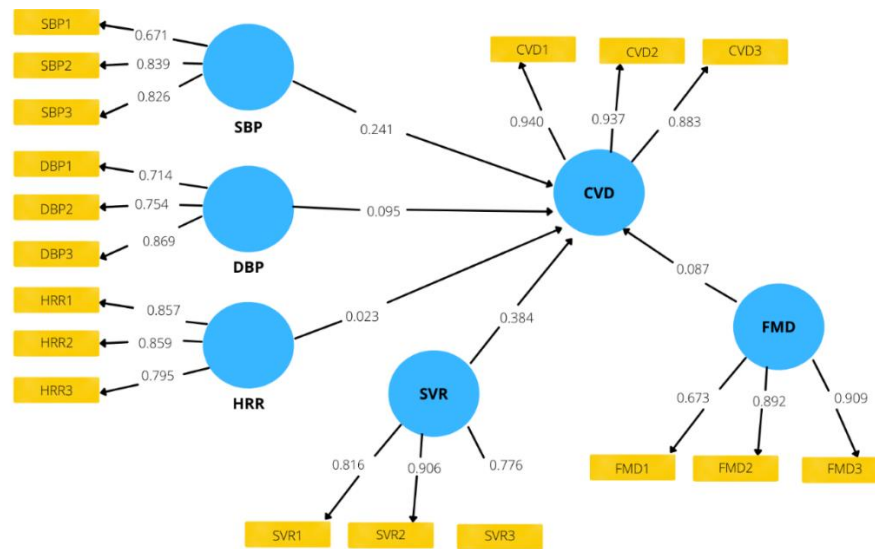


Fig. 3 Path model – hemodynamic variables.

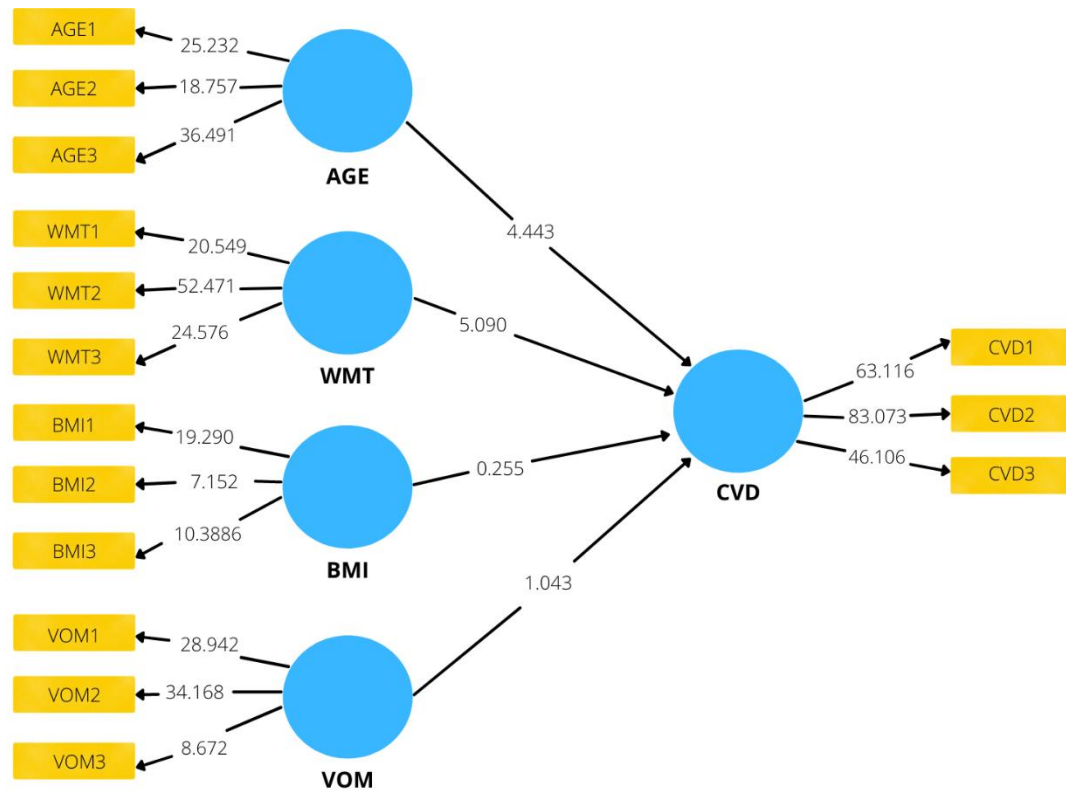


Fig. 4 t-statistic – physical variables.

**4. Discussions**

The hypothesis testing has revealed that among the physical variables under consideration in this research, an individual’s age is significantly associated with the risk of CVD. Among the various risk score measurement instruments, Framingham Risk Score<sup>[47]</sup> is a widely accepted metric of measurement, and it includes age as a primary risk factor associated with CVD. All nine CVD risk assessment metrics considered in this research have considered age to be a factor in assessing the risk of CVD. American Heart Foundation (AHF) has found that the death percentages due to CVD for adults in the age

group of 40-59, 60-79, and above 80 years were approximately 40%, 75%, and 86%, respectively,<sup>[48,49]</sup> and found explicitly that among many other predictors, age has emerged out as a significant predictor of CVD and also that the higher the age, the higher is the risk of CVD. Age is a non-modifiable risk factor of CVD, and the Framingham study indicates that it is mainly associated with other risk factors of CVD that make age-associated risk factors more dominant.<sup>[50]</sup> For instance, for adults less than 50 years of age, smoking doubles CVD’s risk compared to non-smokers, and for adults over 65 years of age, 68% die from CVD if found to be diabetic.<sup>[12]</sup> The deterioration

of cardiovascular functionality with age has been considered the primary cause of CVD.<sup>[51]</sup>

The results have indicated that the weight of an individual is also a significant predictor of CVD among the physical variables. This finding can be corroborated very well by earlier research. Many researchers have associated overweight and CVD risk,<sup>[48,52,53]</sup> It is found that every death out of five in the world is due to overweight or obesity,<sup>[54]</sup> and a meta-analysis of 97 works of literature with a sample size of about 2.9 million people has indicated that obesity is the primary cause of mortality.<sup>[55]</sup> Research has shown that being overweight has many influences on the hemodynamic variables:<sup>[56]</sup>

- increased blood volume increased stroke volume, increased arterial pressure, increased left ventricle (LV) wall stress, pulmonary artery hypertension;
- the cardiac structure, which includes LV concentric

remodeling, LV hypertrophy (eccentric and concentric), left atrial enlargement, and right ventricular hypertrophy;

- the cardiac function, which includes LV diastolic dysfunction, LV systolic dysfunction, and RV failure;
- inflammation, which includes increased C-reactive protein and overexpression of tumor necrosis factor;
- neurohumoral effects include insulin resistance and hyperinsulinemia, leptin insensitivity and hyperleptinemia, reduced adiponectin, sympathetic nervous system activation, activation of the renin-angiotensin-aldosterone system and overexpression of peroxisome proliferator-activator receptor, and cellular effects, which include hypertrophy, apoptosis, and fibrosis.

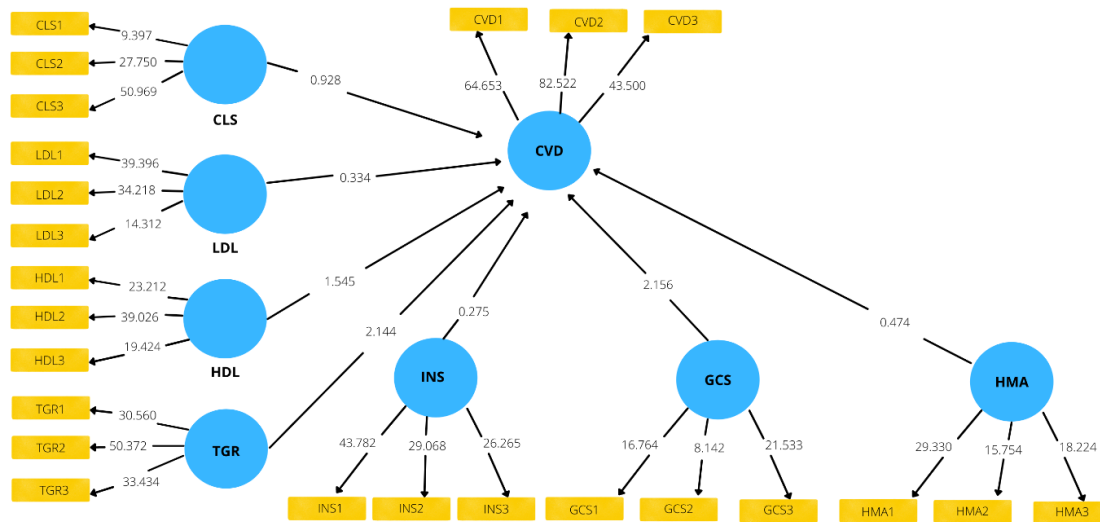


Fig. 5 t-statistic – metabolic variables.

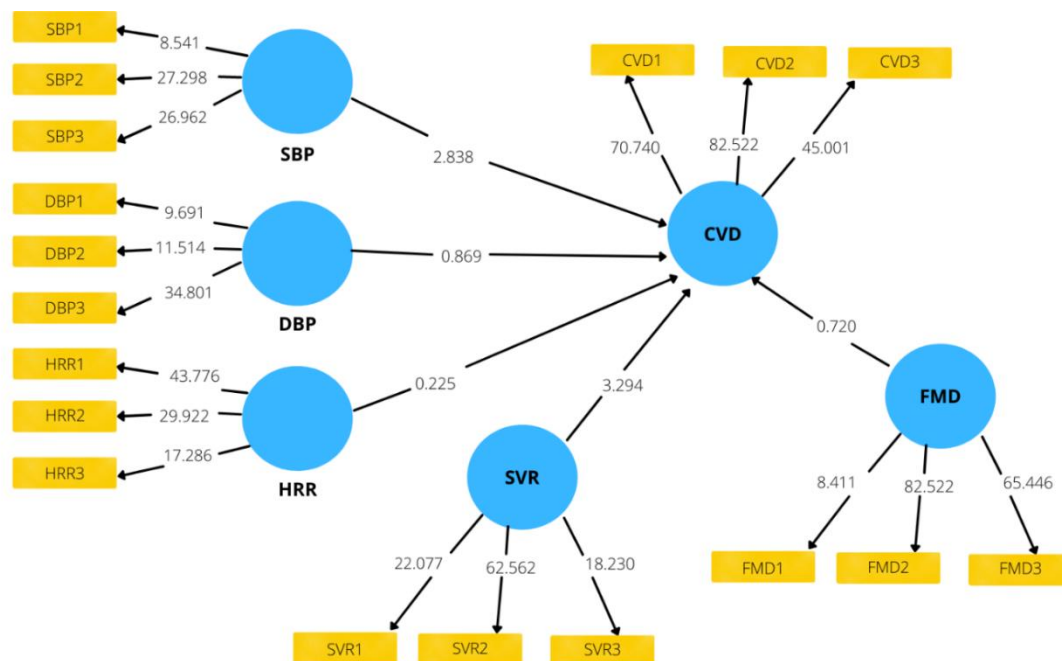


Fig. 6 t-statistic – hemodynamic variables.

All these studies agree with the finding of this research; however, it contradicts the findings of a group of researchers.<sup>[25-27]</sup> While weight has emerged as a predictor of CVD risk, BMI has no significant influence on the risk of CVD, according to the findings of this study. This finding supports the researchers' argument on the 'obesity paradox' who claim that BMI lacks diagnostic performance.<sup>[53,57]</sup> Among the metabolic variables, GCS has emerged as a significant predictor of CVD risk. Jeong *et al.* (2018), with a sample size of 2,682,045 young adults in the age group of 20 to 39 years, proved that high GCS ( $\geq 240$  mg/dL) has a high risk and low GCS ( $<180$  mg/dL) has a low risk of CVD. Whitehall Study with a sample size of 10,025, the Paris Prospective Study with a sample size of 6,629, and also the Helsinki Policeman Study with a sample size of 631, all have found an association of GCS with CVD.<sup>[58]</sup> It was found that there was a 12% increase in CVD risk for an increase of 1 mmol/L increase in fasting glucose.<sup>[59]</sup> It is found that the glucose-CVD relationship is attributed to the direct toxic effect of glucose on cellular function and structure.<sup>[60]</sup> Further, the indirect effects are associated with inadequate insulin secretion and insulin resistance, and hyperinsulinemia before glucose elevations.<sup>[61]</sup>

In the category of metabolic variables, TGR seems to be a significant CVD predictor. The role of TGR on the risk of CVD was found in the early 80s; however, treating healthy people to lower TGR levels was kept under suspension as the evidence then produced was inadequate.<sup>[62]</sup> Research shows that clinical emphasis on reducing CVD risk was primarily on lowering LDL cholesterol; however, through genetic studies when it was found that it was not effective, the focus shifted now to TGR.<sup>[63]</sup> High TGR concentrations lead to remnants rich in cholesterol, and when it enters intima, it can lead to inflammation, foam cell formation, and atherosclerotic plaques, which can cause CVD.<sup>[64,65]</sup> Many researchers have associated TGR with CVD risk.<sup>[66,67]</sup> Researchers have found that TGR is measurable with higher precision, but their concentration varies at a very high rate daily, and that is the reason, statistically speaking, HDL cholesterol has been a better predictor of CVD in the past.<sup>[68,69]</sup>

Among the hemodynamic variables under consideration in this research, SVR has emerged as a significant CVD predictor. It is found that the hyper-adrenergic response, which increases the SVR and the prolonged increase can cause CVD,<sup>[70,71]</sup> have found that the interaction between SVR and ventricular function can cause CVD. Increased SVR results in concentric left ventricular hypertrophy, which can cause CVD.<sup>[72]</sup> It has been found that there are cases where CVD not associated with elevated blood pressure could be due to increased SVR, and hence, it is a significant predictor of CVD.<sup>[73]</sup>

It is revealed in this study that SBP is also a significant predictor of CVD as a hemodynamic variable. Based on the Framingham data, which has been collected over 30 years, it has been found that SBP does have a significant influence on the risk of CVD at all ages in both genders.<sup>[74]</sup> In a study conducted in Taipei City in Taiwan with a sample size of

39,280 men and 38,109 women, it was found that the mortality risks due to CVD were significant for blood pressure values above 160 mm of Hg.<sup>[75]</sup> In another study with a sample size of 1.25 million people, it was found that compared to DBP, the elevated SBP had a stronger influence on CVD; however, surprisingly, this study disregarded the association of blood pressure with CVD risk.<sup>[76]</sup>

#### 4. Conclusion

This research has made a sincere effort in investigating the significance of the association of the antecedents with the risk of CVD. On the overall basis, it can be concluded that even though all the factors considered among physical, metabolic, and hemodynamic variables have no significant association with the risk of CVD, some of the factors of all three variables do have an association. The study illustrates that machine learning does have a role in detecting CVD indicators based on the analysis of data-driven models. It is observed that in the existing literature, the choice of the factors under each of the three variables considered to be having a bearing on the risk of CVD has been by the choice of the individual or group of researchers. Identifying the factors under each of the three variables will be ongoing as newer factors will be found with the research progress. So, data-driven models do have a significant role to play in identifying the most significant predictors among these. Even though this research has specifically narrowed down to age, weight, total glucose, triglycerides, SVR, and SBP as the significant predictors of CVD among the three categories of variables considered, this finding should only be considered as a support in the decision-making of the medical practitioners. More research may be required to confirm the findings in terms of controlled experimentation before using them as the basis of decision-making.

#### 5. Limitations and scope for future work

There have been some limitations in this research, specifically in terms of the sample size and mathematical approximation. According to the principles of SEM, a sample size of above 200 is adequate for analysis, as the data can be extrapolated to a higher number through the bootstrapping technique available in the SmartPLS<sup>®</sup> software. However, extrapolation is still an approximation, and this is a limitation of the study. Further, the SEM analysis is based on the partial least square method, which is also an approximate method. Finally, this research has identified the best predictors of CVD out of the given number of factors considered as most relevant through the existing literature, and hence, there is a possibility that there could be other factors that may influence CVD. Despite these limitations, this research's findings could be of immense help to medical practitioners as a decision tool in CVD risk factors. These limitations provide the scope for future work in this research. The research is based only on three dimensions of possible predictors of CVD, other possible predictors such as lifestyle, family history, working conditions, *et al.*, may be

considered for possible influences on the risk of CVD. Further, the study is limited to the patient data by McGavock, Anderson & Lewanczuk, so other databases can also be explored to make the results more generalizable.

### Acknowledgments

This work was carried out in the Manipal Academy of Higher Education, Manipal, India, and we thank you for providing us with the laboratory facilities.

### Conflict of interest

There are no conflicts to declare.

### Supporting information

Applicable.

### References

- [1] V. B. Kolachalama, P. S. Garg, *Npj Digital Medicine*, 2018, **1**, 54, doi: 10.1038/s41746-018-0061-1.
- [2] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, R. Kim, *JAMA*, 2016, **316**, 2402-2410, doi: 10.1001/jama.2016.17216.
- [3] D.S.W. Ting, C.Y.L. Cheung, G. Lim, G.S. W. Tan, N.D. Quang, A. Gan, A. E.Y.M Wong, *JAMA*, 2017, **318**, 2211-2223, doi: 10.1001/jama.2017.18152.
- [4] S. A. Wartman, C. Donald Combs, *Academic Medicine*, 2018, **93**, 1107-1109, doi: 10.1097/acm.0000000000002044.
- [5] H. S. Wald, P. George, S. P. Reis, J. S. Taylor, *Academic Medicine*, 2014, **89**, 380-386, doi: 10.1097/acm.0000000000000131.
- [6] G.D. Magoulas, A. Prentza, *Springer*, Berlin, Heidelberg, 2001, 300-307, doi: 10.1007/3-540-44673-7\_19.
- [7] S. Saria, A. Butte, A. Sheikh, *PLOS Medicine*, 2018, **15**, 1-5, doi: 10.1371/journal.pmed.1002721.
- [8] A. Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina, T. Botsis, G. Hartvigsen, *Artificial Intelligence in Medicine*, 2019, **98**, 109-134, doi: 10.1016/j.artmed.2019.07.007.
- [9] J. A. Sidey-Gibbons, C. J. Sidey-Gibbons, *BMC Medical Research Methodology*, 2019, **19**, 1-18, doi: 10.1186/s12874-019-0681-4.
- [10] X. Song, A. Mitnitski, J. Cox, K. Rockwood, *MEDINFO*, 2004, **107**, 736-740, doi: 10.3233/978-1-60750-949-3-736.
- [11] M. Fatima, M. Pasha, *Journal of Intelligent Learning Systems and Applications*, 2017, **9**, 1-16, doi: 10.4236/jilsa.2017.91001.
- [12] R. Hajar, *Heart Views*, 2017, **18**, 109-114, doi: 10.4103/HEARTVIEWS.HEARTVIEWS\_106\_17.
- [13] J. M. McGavock, T. J. Asnderson, R. Z. Lewanczuk, *American Journal of Hypertension*, 2006, **19**, 701-707, doi: 10.1016/j.amjhyper.2006.01.013.
- [14] R. K. Simmons, R. L. Coleman, H. C. Price, R. R. Holman, K.-T. Khaw, N. J. Wareham, S. J. Griffin, *Diabetes Care*, 2009, **32**, 708-713, doi: 10.2337/dc08-1918.
- [15] A. Sandoo, N. Chanchlani, J. Hodson, J. P. Smith, K.M. Douglas, G.D. Kitas, *Clinical and Experimental Rheumatology*, 2014, **32**, 914-921, PMID: 25152090.
- [16] R.M. Conroy, K. Pyörälä, A. E. Fitzgerald, S. Sans, A. Menotti, G. Backer, I. Njølstad, *European Heart Journal*, 2003, **24**, 987-1003, doi: 10.1016/s0195-668x(03)00114-3.
- [17] C. J. Lavie, R.V. Milani, H.O. Ventura, *Journal of the American College of Cardiology*, 2009, **53**, 1925-1932, doi: 10.1016/j.jacc.2008.12.068.
- [18] M. Woodward, P. Brindle, H. Tunstall-Pedoe, *Cohort (SHHEC), Heart*, 2005, **93**, 172-176, doi: 10.1136/hrt.2006.108167.
- [19] L. G. H. Goh, S. S. Dhaliwal, T. A. Welborn, A. H. Lee, P. R. Della, *BMJ Open*, 2014, **4**, e004702, doi: 10.1136/bmjopen-2013-004702.
- [20] V. Hainer, I. Aldhoon-Hainerová, *Diabetes Care*, 2013, **36**, S276-S281, doi: 10.2337/dcs13-2023.
- [21] C. McGorrian, S. Yusuf, S. Islam, H. Jung, S. Rangarajan, A. Avezum, A.L. Dans, *European Heart Journal*, 2010, **32**, 581-589, doi: 10.1093/eurheartj/ehq448.
- [22] S. M. Robson, S. C. Couch, J. L. Peugh, K. Glanz, C. Zhou, J. F. Sallis, B. E. Saelens, *Journal of the Academy of Nutrition and Dietetics*, 2016, **116**, 984-990, doi: 10.1016/j.jand.2016.02.011.
- [23] M. Vuolo, J. Staff, *Pediatrics*, 2013, **132**, e568-e577, doi: 10.1542/peds.2013-0067.
- [24] H. C. H Ho, E. Maddaloni, R. Buzzetti, *Diabetes/Metabolism Research and Reviews*, 2019, **35**, doi: 10.1002/dmrr.3134.
- [25] B. Mørkedal, L. J. Vatten, P. R. Romundstad, L. E. Laugsand, I. Janszky, *Journal of the American College of Cardiology*, 2014, **63**, 1071-1078, doi: 10.1016/j.jacc.2013.11.035.
- [26] F. B. Ortega, D.-C. Lee, P. T. Katzmarzyk, J. R. Ruiz, X. Sui, T. S. Church, S. N. Blair, *European Heart Journal*, 2013, **34**, 389-397, doi: 10.1093/eurheartj/ehs174.
- [27] E. Oliveros, V. K. Somers, O. Sochor, K. Goel, F. Lopez-Jimenez, *Progress in Cardiovascular Diseases*, 2014, **56**, 426-433, doi: 10.1016/j.pcad.2013.10.003.
- [28] R.M. Mishra, *SSM - Population Health*, 2019, **7**, 1-7, doi: 10.1016%2Fj.ssmph.2019.100365.
- [29] A. Steptoe, M. Kivimäki, *Annual Review of Public Health*, 2013, **34**, 337-354, doi: 10.1146/annurev-publhealth-031912-114452.
- [30] N. Glozier, G. H. Tofler, D. M. Colquhoun, S. J. Bunker, D. M. Clarke, D. L. Hare, I. B. Hickie, J. Tatoulis, D. R. Thompson, A. Wilson, M. G. Branagan, *Medical Journal of Australia*, 2013, **199**, 179-180, doi: 10.5694/mja13.10440.
- [31] R. S. Vasan, J. M. Massaro, P. W. F. Wilson, S. Seshadri, P. A. Wolf, D. Levy, R. B. D'Agostino, *Circulation*, 2002, **105**, 48-53, doi: 10.1161/hc0102.101774.
- [32] J. Benjamin, M. Blaha, S. E. Chiuve, M. Cushman, S. R. Das, R. Deo, M. C. Jiménez, *Circulation*, 2017, **135**, e146-e603, doi: 10.1161/cir.0000000000000485.
- [33] W. Hinton, A. McGovern, R. Coyle, T. S. Han, P. Sharma, A. Correa, F. Ferreira, S. de Lusignan, *BMJ Open*, 2018, **8**, e020282, doi: 10.1136/bmjopen-2017-020282.
- [34] D. M. Lloyd-Jones, E. P. Leip, M. G. Larson, R. B. D'Agostino, A. Beiser, P. W. F. Wilson, P. A. Wolf, D. Levy,

- Circulation*, 2006, **113**, 791-798, doi: 10.1161/circulationaha.105.548206.
- [35] H. Khan, S. Kunutsor, R. Rauramaa, K. Savonen, A. P. Kalogeropoulos, V. V. Georgiopoulou, J. Butler, J. A. Laukkanen, *European Journal of Heart Failure*, 2014, **16**, 180-188, doi: 10.1111/ejhf.37.
- [36] K. Ichikawa, S. Sakuragi, T. Nishihara, M. Tsuji, A. Mori, F. Yokohama, T. Wada, D. Hasegawa, K. Kawamoto, M. Tanakaya, Y. Katayama, H. Ito, *Heart*, 2018, **104**, 318-323, doi: 10.1136/heartjnl-2017-311751.
- [37] E. Wolsk, R. Bakkestrøm, J. H. Thomsen, L. Balling, M. J. Andersen, J. S. Dahl, C. Hassager, J. E. Møller, F. Gustafsson, *JACC: Heart Failure*, 2017, **5**, 337-346, doi: 10.1016/j.jchf.2016.10.012.
- [38] L. M. Raffield, F.-C. Hsu, A. J. Cox, J. J. Carr, B. I. Freedman, D. W. Bowden, *Diabetology & Metabolic Syndrome*, 2015, **7**, 58, doi: 10.1186/s13098-015-0055-y.
- [39] A. Alavifar, M. Karimimalayer, M.K. Anuar, *Engineering Science and Technology: An International Journal*, 2012, **2**, 326-329.
- [40] J. Holloway, K. Mengersen, *Remote Sensing*, 2018, **10**, 1365, doi: 10.3390/rs10091365.
- [41] J. Fan, K. Wang, Y. Zhong, Z. Zhu, *Statistical Science*, 2021, **36**, 303-327, doi: 10.1214/20-sts785.
- [42] T. N. Beran, C. Violato, *BMC Research Notes*, 2010, **3**, 267, doi: 10.1186/1756-0500-3-267.
- [43] C. Violato, K. G. Hecker, *Teaching and Learning in Medicine*, 2007, **19**, 362-371, doi: 10.1080/10401330701542685.
- [44] K.S. Taber, *Research in Science Education*, 2018, **48**, 1273-1296, doi: 10.1007/s11165-016-9602-2.
- [45] E. E. Rigdon, C.M. Ringle, M. Sarstedt, *Rev's Market Research*, 2010, **7**, 255-296, doi: 10.1108/S1548-6435(2010)0000007011.
- [46] J.F. Hair, G.T.M. Hult, C. Ringle, M. Sarstedt, *International Journal of Research & Method in Education*, 2016, **38**, 220-221, doi: 10.1080/1743727X.2015.1005806.
- [47] R. B. D'Agostino Sr, R. S. Vasan, M. J. Pencina, P. A. Wolf, M. Cobain, J. M. Massaro, W. B. Kannel, *Circulation*, 2008, **117**, 743-753, doi: 10.1161/circulationaha.107.699579.
- [48] A. Yazdanyar, A. B. Newman, *Clinics in Geriatric Medicine*, 2009, **25**, 563-577, doi: 10.1016/j.cger.2009.07.007.
- [49] R. Dhingra, R. S. Vasan, *Medical Clinics of North America*, 2012, **96**, 87-91, doi: 10.1016/j.mcna.2011.11.003.
- [50] W.B. Kannel, R.S. Vasan, *American Journal of Cardiology*, 2009, **104**, 1307-1310, doi: 10.1016/j.amjcard.2009.06.051.
- [51] J. L. Rodgers, J. Jones, S. I. Bolleddu, S. Vanthenapalli, L. E. Rodgers, K. Shah, K. Karia, S. K. Panguluri, *Journal of Cardiovascular Development and Disease*, 2019, **6**, 19, doi: 10.3390/jcdd6020019.
- [52] L.Akil, H.A. Ahmad, *Journal of Health Care for the Poor and Underserved*, 2011, **22**, 61-72, doi: 10.1353/hpu.2011.0166.
- [53] A. De Schutter, C. J. Lavie, D. A. Patel, S. M. Artham, R. V. Milani, *The American Journal of Cardiology*, 2013, **111**, 657-660, doi: 10.1016/j.amjcard.2012.11.013.
- [54] C. J. Lavie, P. A. McAuley, T. S. Church, R.V. Milani, S. N. Blair, *The American Journal of Cardiology*, 2014, **63**, 1345-1354, doi: 10.1016/j.jacc.2014.01.022.
- [55] K. M. Flegal, B. K. Kit, H. Orpana, B. I. Graubard, *JAMA*, 2013, **309**, 71-82, doi: 10.1001/jama.2012.113905.
- [56] C. J. Lavie, M. A. Alpert, R. Arena, M. R. Mehra, R. V. Milani, H. O. Ventura, *JACC: Heart Failure*, 2013, **1**, 93-102, doi: 10.1016/j.jchf.2013.01.006.
- [57] A. Romero-Corral, V. K. Somers, J. Sierra-Johnson, M. D. Jensen, R. J. Thomas, R. W. Squires, T. G. Allison, J. Korinek, F. Lopez-Jimenez, *European Heart Journal*, 2007, **28**, 2087-2093, doi: 10.1093/eurheartj/ehm243.
- [58] J.E. Reusch, C. C. L. Wang, *The Journal of Clinical Endocrinology & Metabolism*, 2011, **96**, 2367-2376, doi: 10.1210/jc.2010-3011.
- [59] N. Sarwar, P. Gao, S. R. Seshasai, R. Gobin, S. Kaptoge, S. J. Danesh, *Lancet*, 2010, **375**, 2215-2222, doi: 10.1016/S0140-6736(10)60484-9.
- [60] S.E. Capes, H.C. Gerstein, *Evid based Cardiology*, 2003, **2**, 161-69, doi: 10.1002/9780470986882.
- [61] M. Stumvoll, B. J. Goldstein, T. W. van Haeften, *The Lancet*, 2005, **365**, 1333-1346, doi: 10.1016/s0140-6736(05)61032-x.
- [62] S. B. Hulley, R. H. Rosenman, R. D. Bawol, R. J. Brand, *New England Journal of Medicine*, 1980, **302**, 1383-1389, doi: 10.1056/nejm198006193022503.
- [63] A. Varbo, M. Benn, A. Tybjaerg-Hansen, A. B. Jørgensen, R. Frikke-Schmidt, B. G. Nordestgaard, *Journal of the American College of Cardiology*, 2013, **61**, 427-436, doi: 10.1016/j.jacc.2012.08.1026.
- [64] I. J. Goldberg, R. H. Eckel, R. McPherson, *Arteriosclerosis, Thrombosis, and Vascular Biology*, 2011, **31**, 1716-1725, doi: 10.1161/atvbaha.111.226100.
- [65] J. C. Rutledge, A. E. Mullick, G. Gardner, I. J. Goldberg, *Circulation Research*, 2000, **86**, 768-773, doi: 10.1161/01.res.86.7.768.
- [66] M. Miller, N.J. Stone, C. Ballantyne, V. Bittner, M.H. Criqui, H.N. Ginsberg, T.A. Lennie, *Circulation*, 2011, **123**, 2292-2333, doi: 10.1161/CIR.0b013e3182160726.
- [67] H. Tada, A. Nohara, M. A. Kawashiri, *Nutrients*, 2018, **10**, 1789, doi: 10.3390/nu10111789.
- [68] A. Langsted, J. J. Freiberg, B. G. Nordestgaard, *Atherosclerosis Supplements*, 2008, **9**, 174, doi: 10.1016/s1567-5688(08)70701-3.
- [69] B. G. Nordestgaard, A. Varbo, *The Lancet*, 2014, **384**, 626-635, doi: 10.1016/s0140-6736(14)61177-6.
- [70] C. Delong, S. Sharma, (2019), Physiology, Peripheral Vascular Resistance. In StatPearls [Internet]. StatPearls Publishing Retrieved on 27/09/2019 from <https://www.ncbi.nlm.nih.gov/books/NBK538308/>.
- [71] C. Eze-Nliam, M. Stevenson, D. McManus, S. Cheng, V. Ramachandran, N. Schiller, N. Parikh, C. Tsao, *Journal of the American College of Cardiology*, 2016, **67**, 1718, doi: 10.1016/s0735-1097(16)31719-3.
- [72] S.R. Daniels, *Journal of the American College of Cardiology*, 2012, **60**, 2651-2652.
- [73] E. V. Bouzinova, O. Wiborg, C. Aalkjaer, V. V. Matchkov,

*Journal of Cardiovascular Pharmacology*, 2015, **65**, 299-307, doi: 10.1097/fjc.000000000000187.

[74] W. B. Kannel, R. S. Vasan, D. Levy, *Hypertension*, 2003, **42**, 453-456, doi: 10.1161/01.hyp.0000093382.69464.c4.

[75] C. Y. Wu, H. Y. Hu, Y. J. Chou, N. Huang, Y. C. Chou, C. P. Li, *Medicine*, 2015, **94**, e2160, doi: 10.1097/md.0000000000002160.

[76] E. Rapsomaniki, A. Timmis, J. George, M. Pujades-Rodriguez, A.D. Shah, S. Denaxas, B. Williams, *Lancet*, 2014, **383**, 1899-1911, doi: 10.1016/S0140-6736(14)60685-1.

### Author Information



**Dr. Dasharathraj K Shetty** is a faculty member of Department of Humanities and Management, Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal. He is an Author, Columnist, Engineer and Social Entrepreneur. Dasharathraj is a B.E. (Computer Science and Engineering) and has three Post-graduation Degrees - MBA (Finance), MPhil (Management) and M.Tech (Computer Science and Engineering). He was awarded a PhD by MAHE, Manipal. He is also a Certified Microsoft Certified Technology Specialist, Dale Carnegie High Impact Teaching Skills, AIMA Certified Management Trainer and RBNQA Examiner.



**Dr. Lewlyn Lester Raj Rodrigues** is Professor in Humanities Management and formerly working with Mechanical and Manufacturing Engineering Department at Manipal Institute of Technology, Manipal, a constituent college of Manipal University, India. He has B.E. (Mechanical Engineering), MBA (HR & Quality), M.Tech. (Production Engineering Systems Technology), and Ph.D. (System Dynamics). His areas of research interest include System Dynamics, Total Quality Management, Innovation/Knowledge/Technology Management, Manufacturing, and Human Resource Management. Dr. Rodrigues has about 175 papers in various journals and conferences, over 750 citations, and is the Reviewer of JKMP and several International conferences and Associate Editor of International Journal of Social Sciences, and Fellow of WBI, Australia. He has successfully guided ten Ph.Ds. and currently guiding 10 Ph.Ds. He has a patent filed and authored two books one in System Dynamics and the other in Service Quality Marketing.



**Dr. Ajith Kumar Shetty** is consultant in Anaesthesia & Critical Care at Sahyadri Narayana, Multispecialty Hospital, Shivamogga, Karnataka. He is M.B.B.S, D.A., M.D. His research area interests include the field of Cardiology and critical health care.



**Dr. Girish Nair** received his Ph.D. in Management in the areas of Financial Accounting and Industrial Economics and Master's Degree in Commerce, International Business and Business Administration. More than 18 years of teaching, research, and administrative experience with several international publications to his credit in the areas of Corporate Finance and Financial Economics and Hospitality Management. Presented his research and acted as a chair for technical sessions at various conferences. He is also an International Examiner for Ph.D. research at various universities in India. He is one of the Principal Investigators on NPRP research projects funded by Qatar National Research Fund

**Publisher's Note:** Engineered Science Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.