Original Research Article

Statistical Analysis of Hypertension Cases in Savelugu Municipality of Ghana

Mr. Iddrisu Musaminu[1] and Dr. Solomon Sarpong[2]

[1]Savelugu Senior High School, Department Of Mathematics, Savelugu – Ghana
[2]university for Development Studies, Department Of Statistics, Navrongo – Ghana

*Corresponding Author
Mr. Iddrisu Musaminu

Abstract: Hypertension affects over one billion people worldwide, is a major cause of deaths and disability in the world, hence this study was undertaken to observe its prevalence and the chances of getting hypertension in the Savelugu Municipality of Ghana. A cross–sectional study was conducted at the Savelugu Municipal Hospital in the month of March 2018 among 600 adults (18 years or above) visiting the out–patients department (OPD) for various ailments. Principal Components Analysis was used to identify five key components in the data. These components were used as covariates to develop the Logistic Regression model. Again, Path Analysis model was used to examine the inter-relationship of the direct and indirect cause and effect of hypertension in the data. The Logistic Regression Analysis shows that increases in number of cigarette sticks smoked, alcohol strength, heart problem, salt intake, canned food intake, milk intake, educational level, BMI and aging are associated with increased likelihood of exhibiting hypertension, whiles, increases in vegetables and fruit intake and physical activities are associated with decreased likelihood of exhibiting hypertension. From the Path Analysis, it was observed that, when all the factors were held constant, the number of cigarette sticks smoked, salt intake, canned food intake and milk intake were found to influence hypertension positively, whiles, vegetables and fruits intake and physical activity had negative influence on hypertension, with heart problem and BMI acting as mediating factors. This study identified several direct and indirect factors that could influence hypertension. These findings are critical for the development of effective interventions and policies to address problems of hypertension.

Keywords: Hypertension, Out Patients Department, Body Mass Index, Inter-relationship, Prevalence, Exogenous variables, Endogenous variables.

INTRODUCTION

Hypertension affects over one billion people worldwide. It affects one in three adults worldwide, (Manandhar and Raman, 2016). In 2000 the prevalence of hypertension was estimated to be 972 million, with the number predicted to rise to 1.5 billion in 2025 (WHO, 2010). It was estimated in 2008 that around 40% of persons 25 or above have hypertension (Manandhar and Raman, 2016). The reported prevalence of hypertension varies around the world, with the rate as low as 5.2% in rural North Indian and as high as 70% in Poland, (Wan *et al.*, 2014). In Nepal, males have slightly high prevalence rate than female with 36% male and 26.2% females having hypertension (Manandhar and Raman, 2016), whiles in Asian-Pacific region, the rates ranges from 5.47% in men and 7.38% in women (Lawest *et al.,* 2004). In Ghana, the prevalence of hypertension in urban population was estimated to be 28.5% (Amoah, 2005).

Hypertension is a major cause of deaths and disability in the world; it is estimated to cause 7.5 million deaths worldwide, about 12.5% of the total of all deaths (Manandhar and Raman, 2016). It is also estimated that about 600 million people with hypertension are at risk of stroke and cardiac failure (WHO, 2000). Data from WHO (2009) shows that death as a result of non-communicable diseases including Hypertension will rise by 17% in the next decade with the highest rise in African to the level of 27%. Hypertension is ranked the third cause of disability worldwide (Ezzati *et al.,* 2002). It is the number one killer disease in Ghana. Doctors at the Korle Bu Teaching Hospital say almost 70 percent of

all deaths at the hospital are caused by hypertensive conditions. The disease affects nearly one out of every five Ghanaian adults, (myjoyonline, 2007). Life style plays an important role in reducing blood pressure and hence reducing the risk of getting hypertension. Some good life style behaviors are, regular exercise, eating healthy diet, reducing sodium in your diet and limit the amount of alcoholic drink, (Mayo Foundation for Medical Education and Research, 1998 - 2016).

Duta and Ray (2012) studied the prevalence of hypertension among the women in rural areas of west Bengal. A population based cross-sectional study was conducted among 1186 women participant age 18 or more years. The study showed that hypertension was seen to increase with age. Other identifiable significant factors are Body Mass Index (BMI), education and family income. Sampatti *et al.* (2013) considered a community-based cross-sectional study to find out the prevalence of hypertension in rural areas of Maharashtra, India. Two independent blood pressure readings were taken in sitting position by visiting each participant at their home. Overall prevalence of hypertension in the study was 7.24%. Logistic regression analysis identified various factors significantly associated with hypertension were age, sex, salt intake, smoking, alcohol consumption and higher economic status. Owusu *et al.* (2013), conducted a study on logistic regression analysis of risk factors associated with hypertension in Kumasi Metropolis, Ghana. Data was gotten from the cardiac clinic Komfo Anokye Teaching Hospital (KATH), Kumasi. The study revealed that alcohol intake, BMI, age, family history of hypertension were all significant factors that can influence hypertension.

Prevention and reducing the effects of hypertension will require concerted efforts from all. It is important to develop best practices for affordable and effective-base program in screening, identifying and treatment of hypertension. Ghana needs to develop policies and implement through multi-sectorial approach involving the Ministries of Heath and other sectors including Education, Agriculture and Finance among others. The role of research cannot be overlooked, since research will help us to determine the risk factors associated with hypertension and to develop a model that can best predict the chances of getting hypertension. Most studies conducted on hypertension considered only the general effect of risk factors on hypertension. This study will look at the inter relationship (cause and effect) among the risk factors, specifically, The main objective of my study is to determine the direct and indirect effect among the risk factors with the aim of predicting the likelihood of getting hypertension.

MATERIAL AND METHODS
Population and Sampling Design
The study was conducted at the Savelugu municipal hospital, Ghana, with a catchment population of over 91,415(PHC, 2000). It is the general hospital in the municipality with over 40 out-patients visit daily. A cross – sectional study was conducted in the month of March 2018 among adults (18 years or above) visiting Out-Patients Department (OPD) of the hospital for various ailments. All patients available at the OPD at the times of visits are randomly interviewed to obtain the sample size of 600 patients. A pre-structured questionnaire was used to collect data on 300 hypertensive Patients and the same number of non hypertensive patients aged 18 years or above visiting the Out-Patient Department of the hospital. However, all patients with severe ailment, life threatening conditions and severe injuries were excluded from the study. Blood pressure readings, weight and other medical information were taken from their medical folder. The list of all respondents and referrals were recorded by the nursing assistant in order to check repetition as a result of multiple visits to the OPD

The Principal Component Model
Let $X'(x_1, x_2 \ldots x_p)$ represent a vector of P original random variables (age, educational level, height, alcohol intake, etc.), $Y'(y_1, y_2 \ldots y_p)$ represent a p linear random variables of X and W' ($w_{11}, w_{12} \ldots w_{pp}$) represent a vector of weights to form the linear combination of the original variables, then the new variables (components) is given by equation (1) and (2):

$$
\begin{aligned}
y_1 &= w_{11}x_1 + w_{12}x_2 + \cdots + w_{1p}x_p \\
y_2 &= w_{21}x_1 + w_{22}x_2 + \cdots + w_{2p}x_p \\
&\vdots \\
y_p &= w_{p1}x_1 + w_{p2}x_2 + \cdots + w_{pp}x_p
\end{aligned}
$$

……………………………….. (1)

$$Y = W'X$$ .... (2)

For easy interpretation, the correlation (loadings) between the principal components and the original random variables is given by equation (3):

$$\rho = w\sqrt{\bar{\lambda}}$$ …………………………………………………………………………….... (3)

Where λ represents the total variance accounted for by the principal components, *W* represents the weights of the principal components and ρ represents the correlation.

The Logistics Regression Model

We define the following dichotomous variable for hypertension as $H = 0$ if non –hypertensive and $H = 1$ if hypertensive. Also, $Y'(y_1, y_2 \ldots y_p)$ represents the set of independent variables (key components). The probability and odds of being hypertensive given y is given by equation (4) and (6) respectively, whiles equation (7) is the logit form:

$$\Pi(y) = P(H = 1 / y) = \frac{e^{(B_0 + B_1 y_1 + \cdots + B_p y_p)}}{1 + e^{(B_0 + B_1 y_1 + \cdots + B_p y_p)}} \quad \ldots\ldots\ldots (4)$$

$$g(y) = \text{odds}(H = 1 / y) = \frac{\Pi(y)}{1 - \Pi(y)} \quad \ldots\ldots\ldots (5)$$

$$= e^{(B_0 + B_1 y_1 + \cdots + B_p y_p)} \quad \ldots\ldots\ldots (6)$$

$$f(y) = \log[g(y)] = B_0 + B_1 y_1 + \cdots + B_p y_p \quad \ldots\ldots\ldots (7)$$

Where:

$\Pi(y)$ refers to the probability of being hypertensive given y

g(y) refers to the odds of being hypertensive given y

f(y) refers to the logit function of some given predictor, y

B_0 is the intercept obtained from the logit function.

B_i (i = 1, 2, ...p) is the logit function coefficient.

Maximum Likelihood Estimation of Parameter

The dependent observations in logistic regression can be considered to be distributed as a Bernoulli distribution, so for the *ith* observation, s_i is distributed as Ber (p_i), where p_i is the probability of being hypertensive;

$$P(H = s_i) = P_i^{s_i}(1 - p_i)^{1 - s_i} \quad \ldots\ldots\ldots (8)$$

$$L = \prod_{i=1}^{n} P_i^{s_i}(1 - p_i)^{1 - s_i} \quad \ldots\ldots\ldots (9)$$

$$= \prod_{i=1}^{n} \left(\frac{e^{BY}}{1 + e^{BY}}\right)^{s_i} \left(\frac{1}{1 + e^{BY}}\right)^{1 - s_i} \quad \ldots\ldots\ldots (10)$$

Where, $By = B_0 + B_1 y_1 + \cdots + B_p y_p$

$$= \left(\frac{e^{By}}{1 + e^{By}}\right)^{\sum_{i=1}^{n} s_i} \left(\frac{1}{1 + e^{By}}\right)^{\sum_{i=1}^{n}(1 - s_i)} \quad \ldots\ldots\ldots (11)$$

$$Log(L) = Z = \log\{(e^{By})^{\sum_{i=1}^{n} s_i}(1 + e^{By})^{-n}\} \quad \ldots\ldots\ldots (12)$$

$$Log(L) = Z = \sum_{i=1}^{n} s_i \log(e^{By}) - n\log(1 + e^{By}) \quad \ldots\ldots\ldots (13)$$

$$Log(L) = Z = \sum_{i=1}^{n} s_i By - n\log(1 + e^{By}) \quad \ldots\ldots\ldots (14)$$

Taking the first differential of equation (14) and equating to zero;

$$\frac{\partial(Z)}{\partial(B)} = \sum_{i=1}^{n} s_i y - n\frac{1}{1 + e^{By}} \cdot \frac{\partial}{\partial B}(1 + e^{By}) = 0 \quad \ldots\ldots\ldots (15)$$

$$= \sum_{i=1}^{n} s_i y - n\frac{e^{By}}{1 + e^{By}} y = 0 \quad \ldots\ldots\ldots (16)$$

$$= sy - npy = 0 \quad \ldots\ldots\ldots (17)$$

$$= y^T(s - \mu) = 0 \quad \ldots\ldots\ldots (18)$$

Where y is the design matrix, s is a column vector of size n of the observed dependent variable and μ is a column vector of size n of $E(s)$.

By taking the second differential of equation (14) we obtained;

$$\frac{\partial^2(Z)}{(\partial(B))^2} = \frac{\partial}{\partial B}\left(\sum_{i=1}^{n} s_i y - \frac{1}{1 + e^{By}} \cdot \frac{\partial}{\partial B}(1 + e^{By})\right) \quad \ldots\ldots\ldots (19)$$

$$= -\frac{ny e^{By} y}{(1 + e^{By})^2} \quad \ldots\ldots\ldots (20)$$

$$= -y^T np(1-p)y = -y^T Wy \quad \ldots\ldots\ldots (21)$$

Where W is a square matrix of order n with var(s) on the diagonal and zero elsewhere.

By Newton-Raphson method $B^t = B^{t-1} + \left[\frac{\partial^2(Z)}{\left(\partial(B^{t-1})\right)^2}\right]^{-1} \frac{\partial(Z)}{\partial(B^{t-1})}$ $\quad \ldots\ldots\ldots (22)$

$$= B^{t-1} + [y^T Wy]^{-1} y^T(s - \mu) \quad \ldots\ldots\ldots (23)$$

Estimate $\hat{B}$ is obtained by numerical estimate by efficient interactive technique on equation (23)

The Path Analysis Model

Under this study, Alcohol intake, number of cigarette sticks, eating habit, physical activity, and family hypertensive are all considered exogenous variables (the variances of this variables are considered to be caused entirely by variables not included in this

model). BMI, heart problems, and hypertension are all endogenous variables in the model (the variances of this variables are considered to be caused in part by other variables in the model). Paths drawn to endogenous variables are unidirectional. The variance in hypertension due to alcohol intake is hypothesized to be the sum of direct effect from alcohol intake and indirect effects from heart problem and extraneous variables (residuals) not included in the model. Similarly, the variance in hypertensive due to physical activity is hypothesized to be the sum of direct effect from physical activity and indirect effects from BMI, heart problem, and extraneous variables (residuals) not included in the model. The path correlation coefficient (beta coefficient) between two variables is represented by P_{ij} where i indicate the effect and j indicate the cause. A positive P_{ij} means increases in the casual variable will lead to increases in the dependent variable and vice versa. The square of the P_{ij} represents the proportion of the affected variable's variance that is caused directly by the casual variable. For easy analysis, we conduct multiple regressions for each endogenous variable (ℓ) from other variables which have direct effect on (ℓ) and solve simultaneously to obtain the beta coefficients. The beta coefficients from these multiple regressions represent the paths weights or direct effects.

Some exogenous variables also exert indirect effect through intermediary variable(s) on endogenous variables. For, example, the indirect effect of number of sticks smoked on hypertension is hypothesized to be the product of the direct effect coming from number of sticks smoked by the direct effect through heart problem. Finally, the total effect of number of sticks smoked on hypertension is the summation of the direct effect of number of sticks smoked and the indirect effect.

RESULTS AND DISCUSSION
The Results of the Principal Component Model
The principal component model is shown in Table 1. It can be seen that 5 components are extracted based on eigenvalue-greater-than-one principle. Only variables with loadings greater than 0.5 are retained. The component 1 loadings for alcohol strength, number of cigarette sticks and heart problem are 0.876, 0.900 and 0.515 respectively. This means that the correlation between component 1 and these variables is strong and positively correlated. In other words alcohol strength, number of cigarette sticks and Heart problem are very important in forming principal component 1. This component is described as alcohol and cigarette component. It can also be seen that component 2 is loaded high on Educational level, salt intake, canned food intake and milk intake (0.624, 0.540, 0.676, and 0.605). This means that these variables are important in forming principal component 2 hence we would described principal component 2 as eating habit 1 component.

Component 3 is highly correlated with fruits intake and vegetables intake with 0.763 and 0.817 loadings respectively. This component can be termed as eating habit 2 component. Also, physical activity (-0.808) is highly negatively correlated with component 4 whiles BMI (0.58) and age (0.697) are positively correlated with component 4. This means that physical activity on one hand move in opposite direction to BMI and age on the other hand, in other words the effect of BMI and age is offset by physical activity. We would refer to component 4 as BMI component.

Finally, component 5 is highly loaded only on family hypertensive status (-0.697). This component is referred to as family hypertensive component.

Table 1: The Principal Component Model

	COMPONENTS				
	1	**2**	**3**	**4**	**5**
Age				0.697	
educational level		0.624			
alcohol strength	0.876				
number of stick	0.900				
heart problem	0.515				
salt intake		0.54			
canned intake		0.676			
fruits intake			0.763		
vegetable intake			0.817		
milk intake		0.605			
BMI				0.582	
physical activity				-0.808	
Family hypertension					-0.697

The Results of Logistic Regression Model
It can be seen from Table 2 that all the components added significantly to the model (p-value < 0.05). This means that changes in the components will result in changes to the probability of being hypertensive. A unit change in alcohol and cigarette component will increase the log of odds by 1.281, that is the odds of being hypertensive for a unit change in alcohol and cigarette component is approximately 3.6 times greater than the original position, that is if the

original chance of getting hypertension is 10 times, then a unit change will increase it to 36 times.

Also a unit change in eating habit 2 component (fruits and vegetables intake) will increase the log of odds by -1.623. Meaning that as eating habit 2 scores increases the log of odds reduces. In terms of odds, a unit change in eating habit 2 will result in 0.197 times of getting hypertension, in other words if there is originally 100 chances of getting hypertension a unit change in this component will reduce the chances to approximately 20 times.

BMI component has the highest log of odds of 2.082 that is a unit change in BMI component will result in odds of being hypertensive 8 times greater than the original position. The odd ratio of BMI component to alcohol and cigarette component is 2.21, meaning that if there is a 100% chance of getting hypertension through alcohol and cigarette component they will be a 221% chances of getting same through BMI component.

Table 2: The Logistic Regression Model

Key components	B	SE	Wald	df	Sig	Exp(B)	95% C.I for Exp(B)	
							Lower	Upper
alcohol and cigarette	1.281	0.27	26.41	1	0.00	3.621	2.209	5.87
eating habit 1	1.084	0.14	62.229	1	0.00	2.958	2.259	3.872
eating habit 2	-1.623	0.16	101.633	1	0.00	0.197	0.144	0.271
BMI	2.082	0.2	109.778	1	0.00	8.009	5.432	11.837
family hypertensive	0.415	0.15	8.15	1	0.00	1.514	1.139	2.013

Predicted Probabilities and Components Scores for 5 selected Respondents

The component scores are indices of risk factors. The higher the score the higher the chances of getting hypertension. A negative score can be considered as a safety factor. Table 3 is an extract of predicted probabilities of being hypertensive of the 600 sampled respondents, it can be seen from the Table that the chances that the respondent 6 is hypertensive is approximately 1%, that is if 100 persons of similar characteristics are sampled only one is expected to be hypertensive. Since the respondent score very low (-2.03) on BMI component as compared to the other components, we can conclude that this may be due to the fact that the respondent engaged in physical activity (leisure activities and household activities) which offsets the influence of BMI and aging. However, respondent 87 score high (1.34) on BMI components but this time the chances that the respondent is

hypertensive is 92%, this can be attributed to high BMI or aging. It can also be seen that respondent 51 score high (5.20) on alcohol and cigarette component as compared to the other component. Since this respondent is hypertensive (98%), we can say that his predicament is due largely to alcohol intake and/or number of cigarette stick smoked which could affect the heart, hence being hypertensive.

Furthermore, respondent 56 has 98% chance of getting hypertension and since the respondent score high (1.8) on eating habit 1component, it can be attributed to much salt, canned food intake and a lot of milk consumption with high level of education. On the other hand, respondent 33 score very low (-1.09) on eating habit 2 components, this means that the respondent takes a lot of fruits and vegetables and hence non-hypertensive.

Table 3: Selected Component Scores and Predicted Probabilities

Respondent	Alcohol and cigarette component	Eating Habit 1 Component	Eating Habit 2 Component	BMI Component	Family hypertensive Component	Probability
6	-0.34	0.82	0.61	-2.03	0.02	0.01
33	-0.38	-0.25	-1.09	-0.47	-0.2	0.48
51	5.2	-1.43	-0.65	-0.86	-1.4	0.98
56	-0.53	1.8	-1.47	0.17	0.17	0.98
87	0.14	-0.34	-0.07	1.34	-0.49	0.92

Assessing the Distribution of the Residual

Examining the error plots of standardized residual against some linear predators, we can see that 95% of the cases lie within ±3 band with few outliers

for all the plots (Fig. 2 and Fig. 3). This means that only a few individual cases fit the model poorly, hence the model is appropriate and valid.
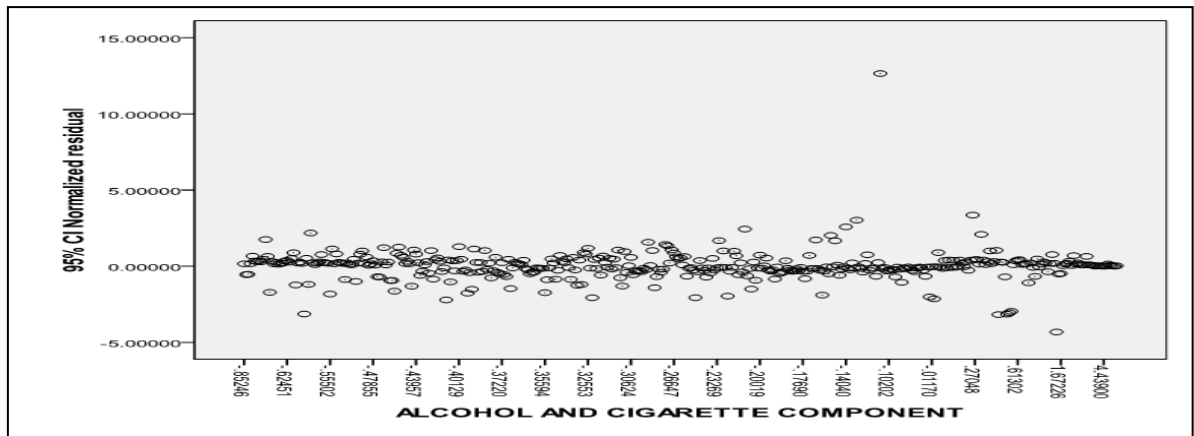
Figure 2: Error plots of standardize residuals against alcohol and cigarette component
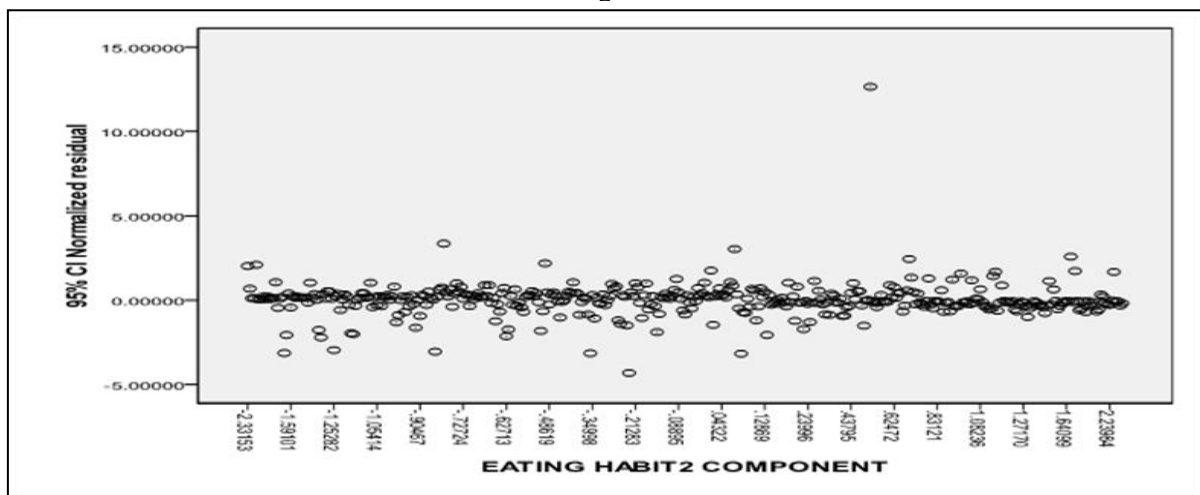

Figure 3: Error plot of standardize residual against eating habit 2 component

The model coefficient of determination (R^2) shows how much change in the dependent variable (hypertension) that can be explained by the model. For our model 49.9% (Cox & Snell R Square) to 66.5% (Negelkerke R Square) of changes in hypertension can be explained by the model. It can also be seen from Table 4 that 250 respondents were classified correctly as being non-hypertensive, representing 83.3%. Similarly, 245 respondents were classified correctly of being hypertensive, representing 81.70%. The overall correct classification stands at 82.5%.

Table 4: Prediction Classification

		Predicted		Percentage
		Non Hypertensive	Hypertensive	
Observed	Non Hypertensive	250	50	83.3
	Hypertensive	55	245	81.7
Overall %				82.5

Results of the Path Model

From Figure 4, it can be seen that the variables alcohol strength, number of cigarette sticks, eating habit 1 (salt intake, canned food intake and milk intake), eating habit 2 (vegetables and fruits intake), physical activity (leisure activity and household activity) and family hypertensive are all exogenous variables whiles heart problem, BMI and hypertension are endogenous variables. The variation in the data attributable to the exogenous variable ranges from 0.3 to 2.8 whiles variation in the data accounted for by extraneous variables (residuals) not included in this model ranges from 0.2 to 10.

The result indicates that hypertension is directly affected by number of cigarette stick smoked, eating habit 1, eating habit 2, physical activity, family hypertensive status, heart problem and BMI, however, surprisingly the direct effect of alcohol intake on hypertension is non-significant. It can also be seen that hypertension is more affected directly and positively by BMI (Std beta weight = 0.246) as compared to the

remaining causes. In other words increases in BMI whiles all other factors are held constant will result in greater impact on hypertension,

However, hypertension is directly affected more negatively by eating habit 2 (Std beta weight = -0.294). That is increases in eating vegetables and fruits intake whiles all other factors are held constant will reduce the chances of getting hypertension faster as compared to physical activity (Std beta weight = -0.156). We can also see that BMI is reduce directly by physical activity (Std beta weigh = -0.300), weakly by eating habit 2 (Std beta weight = -0.151) and cause directly by eating habit 1 (Std beta weight = 0.117).

Finally, heart problem can be reduced by eating habit 2 (Std beta weight = -0.165) but caused by number of cigarette sticks smoked (Beta weight= 0.276). Surprisingly, the direct effect of family hypertensive statues on heart problem is negative.

Table 5 also shows the indirect effects in the model. It can be observed that hypertension is significantly (P-value < 0.01) affected indirectly by physical activity, eating habit 2, eating habit 1 and number of cigarette sticks smoked with the effects ranging from –0.056 to 0.031. It can also be seen that the indirect effects of physical activity and eating habit 2 (without holding the effects of heart problem and BMI constant) are both negative indicating that increases in physical activity and eating habit 2 will lead indirectly to reduction in hypertension as compared to the indirect effects of number of cigarette sticks and eating habit 1 (salt intake, canned food and milk intake) which are both positive. Overall, BMI (0.246) has the highest causing factor of hypertension and eating habit 2 (-0.349) can substantively reduce hypertension (p–values< 0.01) as shown in Table 6.
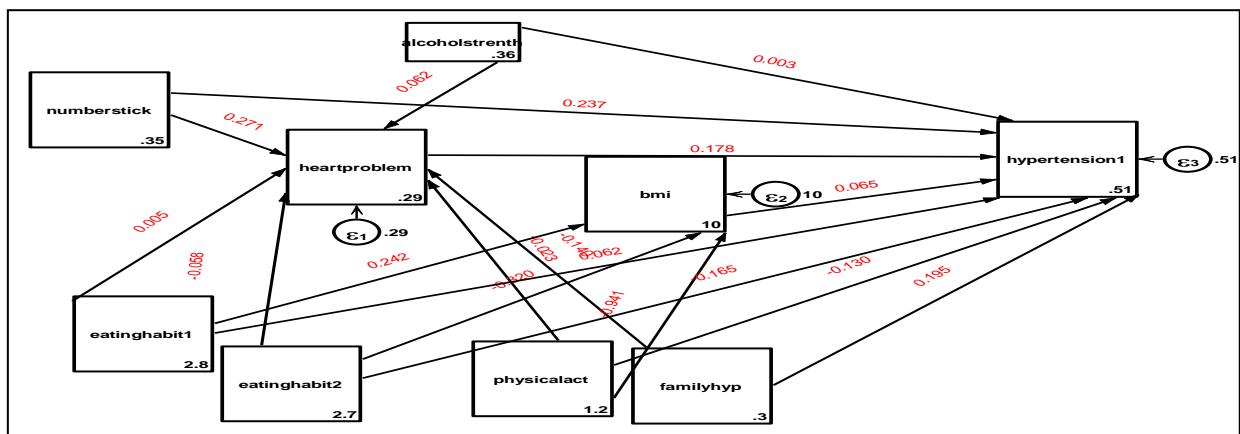

Figure 4: The Final Model

Table 5: Indirect Effects of Variables

	Coefficient	Std. Coefficient	$P> |z|$
hypertension ←			
alcohol strength	0.011	0.007	0.259
number stick	0.048	0.031	0.005
eating habit 1	0.017	0.03	0.008
eating habit 2	-0.031	-0.055	0.000
Physical act.	-0.066	-0.079	0.000
Family hyp	-0.024	-0.015	0.018

Table 6: Total Effects of Variables

	Coefficient	Std. Coefficient	$P> |z|$
hypertension ←			
heart problem	0.178	0.113	0.001
BMI	0.065	0.246	0.000
alcohol strength	0.015	0.009	0.834
number stick	0.285	0.183	0.000
eating habit 1	0.079	0.143	0.000
eating habit 2	-0.196	-0.349	0.000
Physical act.	-0.196	-0.235	0.000
Family hyp	0.170	0.100	0.002

Model Fit Measures (Final Model)

The chi-square value of 5.631 (p-value = 0.228) is non-significant at the 0.05 level. This result

Suggests that the final model fits the data well and the RMSEA corroborate this evidence with fit statistic of 0.026 well below the recommended 0.06 cutoff. Again,

the Tucker-Lewis Index of 0.983 is considerable above the 0.95 threshold indicating a satisfactory model fit (Table 7).

Table 7: Model Fit Measures (Final Model)

Fit statistic	Value
Chi-square	5.631
$P>$ Chi-square	0.228
RMSEA	0.026
TLI	0.983

DISCUSSIONS

Fourteen variables including age, educational level, alcohol intake, number of cigarette sticks smoked, heart problem, salt intake, canned food intake, milk intake, sugar intake, vegetable intake, fruits intake, BMI, physical activity and family hypertensive status were subjected to principal component analysis. Out of the fourteen components formed, only the first five components displayed eigenvalues-greater-than-one, therefore only the first five components accounting for approximately 60% variance in the data were retained. Component 1 was highly loaded on number of cigarette sticks smoked, alcohol intake and heart problem which was subsequently labeled the alcohol and cigarette stick component. Also, four variables, that is educational level, salt intake, canned food intake and milk intake were loaded highly on the second component which was labeled eating habit 1 component. The rest of the components are eating habit 2 component, BMI component and family hypertensive component. The component scores which are indices of risk factors are used for further regression analysis.

A logistic regression analysis was performed to determine the influence of key comportments on the likelihood that the respondents have hypertension. The chi-square value was 414.20 (p-value < 0.01), meaning that the model was statistically significant. Also, the model has good predictive power and thus explained 66.5% (Nagelkerke R^2) of the variance in hypertension and correctly classified 82.5% of cases. Increasing the alcohol and cigarette component, eating habit 1 component, BMI component and family hypertensive component were associated with increased likelihood of exhibiting hypertension, but increasing in eating habit 2 was associated with reducing hypertension. Diagnoses of hypertension normally requires several readings of high blood pressure, however, with this model we can easily predict the likelihood of getting hypertension, hence proper measures to prevent or reduce the effect of hypertension can be taken. Again, the model can predict the nature of the risk factors of hypertension for an individual, for instant the nature of risk factor for respondent 51 is due to smoking, alcohol consumption and heart problem (alcohol and number of cigarette stick component) and since the choice of treatment depends on the individual conditions, doctors can easily prescribe the best treatment for the respondent.

Finally, from the path analysis model, the results identified several direct and indirect causes of hypertension, specifically, number of cigarette stick smoked, eating habit 1, eating habit 2, physical activity, family hypertension, heart problem and BMI. The study identifies number of cigarette smoked as the strongest significant factor, directly and indirectly affecting the heart and consequently causing hypertension. This is an empirical confirmation of a well-known theory that smoking carries nicotine and carbon monoxide which can cause damage to the blood vessels and the heart and hence causing hypertension (Yuanxin, 2013). The results convey an important message to interventionist that for a person who has heart problem and smokes, solving the heart problem without addressing his/her smoking habit will probably not produce a significantly change in hypertension. Also, eating habit 1 (salt, canned food and milk intake) is directly related to hypertension and indirectly related to hypertension through mediating factors such as BMI. Again, this support the theory that too much salt and high level of cholesterol may not keep the blood healthy and can lead to shrinkage of the artery and thus causing hypertension (Yuanxin, 2013). Any effort aimed at reducing hypertension by reducing eating habit 1 without addressing the mediating factor (BMI) may not produce a significant change in hypertension. The study also supports a known relevance of eating habit 2 (fruits and vegetables intake) and physical activity for good health and checking BMI and thereby reducing hypertension. Moreover, family hypertensive status have direct relationship to hypertension, this could be due to genetic mutation risk of the respondent (based on family hypertensive status). However it's a big surprise that the data did not support the effect of alcohol intake on both the heart and hypertension. The result suggests that family hypertensive status is negatively related to heart problem. Overall the final model fits the data well.

CONCLUSION

Hypertension is one of the public health challenges worldwide. The effects of hypertension can be very devastating and can lead to aneurysm, kidney failure, stroke etc, meanwhile there are no clear signs associated with it until it results in complications and even deaths.

This study considered the application of logistic regression using key components from PCA as predators to predict the chances of getting hypertension. The study also hypothesized a path model to determine the inter-relationship (cause and effect) of risk factors on hypertension. All the models developed were tested and found to be statistically significant and fitted the data very well.

The results identifies five key components that have influence on hypertension, that is alcohol and cigarette component, eating habit 1 component (salt intake, canned food intake, milk intake and educational

level), eating habit 2 component (vegetables and fruits intake), BMI component (BMI, age and physical activity) and family hypertensive component. These components were found to have high predictive power in predicting hypertension and therefore were set as covariates to model the logistic regression. Increasing the alcohol and cigarette component, eating habit 1 component, BMI component and family hypertensive component were associated with increased likelihood of exhibiting hypertension, but increasing the eating habit 2 was associated with reducing hypertension. The result also shows several direct and indirect causes of hypertension, specifically, number of cigarette sticks smoked, eating habit 1 were found to directly/indirectly influence hypertension positively and eating habit 2, physical activity have negative influence on hypertension with heart problem and BMI acting as mediating factors. Surprisingly, family hypertensive statues have negative effect on heart problem and alcohol intake seems not to have significant impact on hypertension. The results of this study are critical to government and other cooperate bodies in developing effective intervention and policies to address problems of hypertension. This study also establishes sound foundation for further investigation of the variables under discussions, perhaps with the use of longitudinal design and survival analysis.

REFERENCES

1. N Manandar, S. Raman. (2016). Survey on the risk factors of hypertension using logistic regression: SCIREA *Journal of Health.*
2. World Health Organization. (WHO, 2010). Health Systems Financing: The Path to Universal Coverage
3. Wan. (2014). Prevalence, Awareness, Treatment, and Control of Hypertension and Associated Risk Factors among Adults in China, *PMC Journal*, 95(14).
4. Lawest. (2004). Hypertension: its Prevalence and Population Attributable Fraction for Mortality from Cardiovascular Disease in the Asian-Pacific Region. *Journal of Human Hypertension,* 25(1), 73-9.
5. Amoah. (2005). The Changing Patterns of Hypertension in Ghana: A Study of four Rural Communities in the GA District.
6. World Health Organization. (WHO, 2010). Health Systems Financing: The Path to Universal Coverage
7. World Health Organization. (WHO, 2009). Cardiovascular disease and Hypertension in Sub-Saharan African.
8. Ezzati,(2002). Selected Major Risk Factors and Global and Regional Burden of Disease. *National center for Biotechnology Information, US National Library of Medicine.*
9. Hypertension epidemic: silent global killer – myjoyonline (2007)
10. Mayo Foundation for Medical Education and Research. (1998 – 2016). High blood pressure (hypotension) – 10 Ways to Control Blood Pressure without Medication, MayoClinic.com.
11. Dutta. & Ray. (2012). The Prevalence of Hypertension among the Women in Rural Areas of West Bengal. *International Society of Hypertension*, 20(4), 219-25.
12. Sampatti. (2013). Community-based Cross-sectional Study on Prevalence of Hypertension in Rural Areas of Maharashtra, Indian. *International Journal of Medicine and Public Health,* 2, pp. 2.
13. Owusu. (2013). Logistic Regression Analysis of Risk Factors Associated with Hypertension in Kumasi, Metropolis, Ghana. *Open Science Journal of Clinical Medicine,*3(6), pp. 205-211.
14. Yuanxin, (2013). Survival Analysis of Cardiovascular Disease. All Theses and Dissertations (ETDs) 1209.