Predictors of High Blood Pressure in South African Children: Quantile Regression Approach

Lyness Matizirofa[1,*] and Anesu Gelfand Kuhudzai[2]

[1]*Department of Statistics, University of Johannesburg, South Africa*

[2]*Statistical Consultation Services, University of Johannesburg, South Africa*

Abstract: *Objective*: To identify predictors of blood pressure (BP) in children and explore the predictors` effects on the conditional quantile functions of systolic blood pressure and diastolic blood pressure.

Methods: A secondary data analysis was performed using data from the South African National Income Dynamics Study (2014-2015). From this particular secondary data, data for children aged between 10-17 years were extracted for analysis. The variables used in the study were systolic blood pressure (SBP), diastolic blood pressure (DBP), body mass index (BMI), age, smoking, alcohol consumption, exercises, gender and race. Two parameter estimation methods were used, ordinary least squares (OLS) and quantile regression (QR).

Results: BMI had positive statistically significant estimated OLS and conditional quantile functions with both the BP measures except the 95th quantile for SBP. Age had also positive statistically significant estimated OLS and QR coefficients except for the 95th percentile, with both DBP and SBP respectively. Gender was found to be inversely related to both DBP and SBP except the 10th quantile for DBP. Race was partially significant to DBP. Smoking, alcohol consumption and exercises did not present any statistically significant relations with both DBP and SBP for all the estimated OLS and QR coefficients.

Conclusion: BMI, age, gender and partially race were found to be predictors of BP in South African children using both OLS and QR techniques. Exercises, smoking and alcohol consumption did not present any statistically significant relations with both DBP and SBP probably because few participants exercise regularly, smoke and drink alcohol to bring out a significant change in both BP measurements.

BACKGROUND

Globally, cardiovascular diseases are major causes of death among general population [1]. Similarly, it has been found in another study that high blood pressure is a major health problem which is associated with increased risk of cardiovascular diseases [2]. Besides being a leading burden of disease worldwide in adults, blood pressure is becoming a matter of concern in children as well [3]. Childhood hypertension has been widely investigated in the last decade [4]. According to Aounallah-Skhiri *et al.* [2] the prevalence of high blood pressure in children is a predictor of hypertension later in life. Numerous studies have shown that the underlying condition of hypertension in adults may have its origin in childhood or adolescence. This was also noted by Ayatollahi *et al.* [5] who reported that the prevalence of high blood pressure in children is ranging from 5.4% to 19.4%.

Several studies have established the risk factors of blood pressure in children as, age, gender, ethnicity, obesity, genetic, environmental, inactive lifestyle, low quality food, smoking habits, behavioural factors and socio-economic status [2, 4-8]. In South Africa hypertension is regarded as one of the greatest health challenges apart from HIV/AIDS [9]. It is also the most prevalent cardiovascular risk factor and the leading contributor to cardiovascular morbidity and mortality [9].

There is no doubt about the burden of blood pressure worldwide. To date, very few studies have identified predictors of childhood blood pressure disease in South Africa. Hence, it was ideal to conduct a secondary data analysis on blood pressure data collected from South African National Income Dynamics Survey which was carried out in 2014-2015. The main aim of this study was to identify predictors of blood pressure in South African children and explore these predictors` effects on the conditional quantile functions of both SBP and DBP.

Data

This study used secondary data obtained from South African National Income Dynamics Survey (NIDS) which was carried out in 2014-2015. The South African National Income Dynamics Survey was embarked by the South African Presidency in order to

*Address correspondence to this author at the Department of Statistics, University of Johannesburg, Doornfontein Campus, P.O. Box 17011, Doornfontein, 2028 Johannesburg, South Africa; Tel: +27 11 559 6648; E-mail: lynessm@uj.ac.za

track changes in the well-being of South Africans [12]. The NIDS 2014-2015 covered the entire country of South Africa of which the lowest level of geographic aggregation covered was district municipality [12].

Ethical approval for the South African National Income Dynamics Survey was granted by the University of Cape Town`s Faculty of Commerce Ethics Committee. From this particular secondary data, data for children aged between 10 – 17 years were extracted for analysis.

Prior to performing data analysis, data cleaning was conducted. The data cleaning process involved dropping 597 observations or cases with missing data for any of the variables used in the study and a representative sample of 2 428 was obtained. Cases with missing data pose a challenge because typical modelling procedures simply discard these cases from the analysis [14]. Expectation maximisation technique could not be used to replace missing data since the missing data were not less than 5% [14].

The variables used in this study are SBP, DBP, age, BMI, smoking, alcohol consumption, exercises, gender and race. The dependent variables were DBP and SBP whilst the independent variables were age, BMI, smoking, alcohol consumption, exercises, gender and race. BMI was calculated as weight (kg)/height (m)^2 and age was calculated as *date of interview - date of birth*. The categorical variables were coded as follows: smoking (1 = Yes and 2 = No), alcohol consumption (1 = I have never drank alcohol, 2 = I no longer drink alcohol, 3 = I drink very rarely, 4 = Less than once a week, 5 = On 1 or 2 days a week, 6 = On 3 or 4 days a week and 7 = On 5 or 6 days a week), Exercises (1 = Never/Less than once a week, 2 = Once a week, 3 = Twice a week and 4 = Three or more times a week), Gender (1 = Male and 2 = Female) and Race (1 = African, 2 = Coloured, 3 = Asian/Indian and 4 = White).

Estimation Methods

In this paper, two parameter estimation methods were used. Thus, the ordinary least squares and quantile regression techniques. Quantile regression is a statistical technique which commonly use the median regression and estimate conditional quantiles whilst ordinary least squares (OLS) method focus on the conditional mean [10]. Beyerlein *et al.* [11] stated that the approach and interpretation of quantile regression are similar to those of linear regression.

However, quantile regression leads to more comprehensive results because of its ability to assess any part of the outcome distribution, whereas linear

regression model can model only the mean of the outcome. Also, quantile regression is more preferable to OLS when using data with extreme values.

According to Leibbrandt *et al.* [12] the quantile regression model for the θ th quantile with $0 \le \theta \le 1$ can be written as:

$$Q_\theta(y \mid x) = x \beta_\theta$$

where vector x contains the set of explanatory variables. β_θ is the vector of the regression parameters (i.e. intercept and regression coefficients). $Q_\theta(y \mid x)$ denotes the θ th quantile of the dependent variable y conditional on the set of explanatory variables contained in the vector x.

Data Analysis

For analysis, IBM SPSS Version 24 was used to conduct frequencies on demographic and life style characteristics of participants and summary statistics on continuous variables such as SBP, DBP, BMI and age. The 10th, 25th, 50th, 75th, and 95th quantile regression coefficient estimates for the SBP and DBP respectively were also produced with the aid of R statistical package namely quantreg written by Koenker [13]. The specified quantiles selected for the quantile regression analysis were derived from Koenker [13].

RESULTS

Table **1** reveals that 1 175 (48.4%) of the participants were males and 1 253 (51.6%) were females. Majority of the participants were African and they were 2 198 (90.5%) and the least number of participants were Asian/Indian and they were 15 (0.6%).

The age ranges from 10 years to 17 years, with 14 years having the most participants 547 (22.5%) and 10 years the least 1 (0.04%). It is evident that the majority of the participants1 697 (69.9%) do not exercise whilst 349 (14.4%) do exercise three or more times a week. Majority of the participants did not smoke and they were 2 327 (95.8%) whilst 101 (4.2%) smoked.

Study findings indicates that 2 214 (91.2%) of the participants never drank alcohol. Only 1 (0.04%) person indicated that he/she drank alcohol 5 or 6 days a week.

Table **2** indicates that the mean BMI for boys and girls were 21.68 kg/m^2 and 23.96 kg/m^2 respectively.

Table 1:   Demographic and Lifestyle Characteristics

CHARACTERISTIC		n	%
Gender	Male	1175	48.4
	Female	1253	51.6
Age	10 years	1	0.04
	12 years	7	0.3
	13 years	379	15.6
	14 years	547	22.5
	15 years	500	20.6
	16 years	489	20.1
	17 years	505	20.8
Race	African	2198	90.5
	Coloured	195	8.0
	Asian/Indian	15	0.6
	White	20	0.8
Exercises	Never/ Less than once a week	1697	69.9
	Once a week	191	7.9
	Twice a week	191	7.9
	Three or more times a week	349	14.4
Smoking	Yes	101	4.2
	No	2327	95.8
Alcohol Consumption	I have never drank alcohol	2214	91.2
	I no longer drink alcohol	49	2.0
	I drink very rarely	117	4.8
	Less than once a week	24	1.0
	On 1 or 2 days a week	19	0.8
	On 3 or 4 days a week	4	0.2
	On 5 or 6 days a week	1	0.04

Table 2:   Summary Statistics of Continuous Variables

	SBP (mmHg)		DBP (mmHg)		Age (Years)		BMI (kg/m^2)	
	Boys	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**
Mean	115.92	110.65	73.26	72.85	15.08	15.06	21.68	23.96
Standard Error	0.44	0.38	0.36	0.33	0.04	0.04	0.13	0.14
Median	115.00	110.00	72.00	72.00	15.00	15.00	20.76	23.14
Standard Deviation	15.00	13.48	12.27	11.72	1.37	1.40	4.47	5.02
Minimum	56	73	30	37	12	10	11.10	11.87
Maximum	193	200	139	132	17	17	53.51	46.61
Range	137	127	109	95	5	7	42.41	34.74
Interquartile Range	19	18	14	15	2	2	4.21	5.61
Skewness	0.59	0.75	0.46	0.55	0.01	-0.10	2.15	1.17

Boys had mean SBP and mean DBP of 115.92 mmHg and 73.26 mmHg respectively whilst girls had mean SBP and mean DBP of 110.65 mmHg and 72.85 mmHg respectively.

Tables **3** and **4** illustrates the results of QR analysis and OLS regression analysis for DBP and SBP as response variables respectively against age, BMI, smoking, alcohol consumption, exercises, gender and race as explanatory variables.

In OLS regression analysis, respondents` age ($p<0.01$), BMI ($p<0.01$), gender ($p<0.01$) and race ($p<0.05$) are statistically significant risk factors for DBP. Smoking, alcohol consumption and exercises are not statistically significant with DBP across all quantiles. BMI is the only statistically significant variable across all quantiles at 1% level of significance. Age and gender do present statistically significant relations with DBP across all other quantiles except 95% and 10% quantiles respectively. Finally, race is significant only at 95% quantile ($p<0.01$).

Figure **1** below presents a concise visual summary of the estimated quantile regression results. Each plot illustrates one of the explanatory variables in the quantile regression model. The dash-dot-dot curve with filled dots represents the 19 point estimates of the coefficient for τ`s ranging from 0.05 to 0.95 [13]. In this regard τ is representing different quantiles. The shaded grey area shows a 95 percent pointwise confidence band. Superimposed on the plot is a solid red line representing the ordinary least squares of the mean, with two dashed red lines representing a 95 percent confidence interval for a particular coefficient. Each plot is showing different quantiles on the x axis and the covariate effect on the y axis. Results illustrated in each plot are similar to those shown in Table **3**.

It is evident from Table **4** that gender ($p<0.01$) is the only variable that is statistically significant with SBP in all quantiles (10%, 25%, 50%, 75% and 95%) whilst smoking, alcohol consumption, exercises and race are not statistically significant with SBP. Age and BMI

Table 3: OLS and QR Coefficients for Diastolic Blood Pressure`s Risk Factors

Variable	Quantile regression results					OLS
	Q10	Q25	Q50	Q75	Q95	
Age	0.91***	0.85***	0.68***	0.57**	0.34	0.61***
BMI	0.34***	0.37***	0.45***	0.49***	0.49***	0.38***
Smoking	-1.12	-1.00	1.84	-2.82	0.61	-0.72
Alcohol Consumption	-0.98	0.03	0.60	-0.42	0.48	0.03
Exercises	-0.14	-0.22	-0.32	-0.39	-0.42	-0.23
Gender	-0.65	-1.43**	-1.98***	-2.76***	-3.02**	-1.49***
Race	1.17	0.71	1.27	0.96	4.52***	1.39**

Note: OLS=Ordinary Least Squares; BMI= Body Mass Index.
$p^{**} < 0.05$ $p^{***} < 0.01$.

Table 4: OLS and QR Coefficients for Systolic Blood Pressure`s Risk Factors

Variable	Quantile regression results					OLS
	Q10	Q25	Q50	Q75	Q95	
Age	0.79***	1.29***	1.09***	1.31***	0.92	1.11***
BMI	0.25***	0.21**	0.32***	0.28***	0.18	0.24***
Smoking	-0.59	-0.05	-1.14	-0.73	-1.21	-1.34
Alcohol Consumption	-0.65	0.32	-0.44	0.36	1.15	0.18
Exercises	-0.07	-0.14	-0.00	-0.08	-0.59	-0.19
Gender	-4.08***	-4.83***	-5.85***	-6.53***	-8.31***	-5.92***
Race	0.93	0.95	0.67	1.46	2.71	1.38

Note: OLS=ordinary least squares; BMI=body mass index.
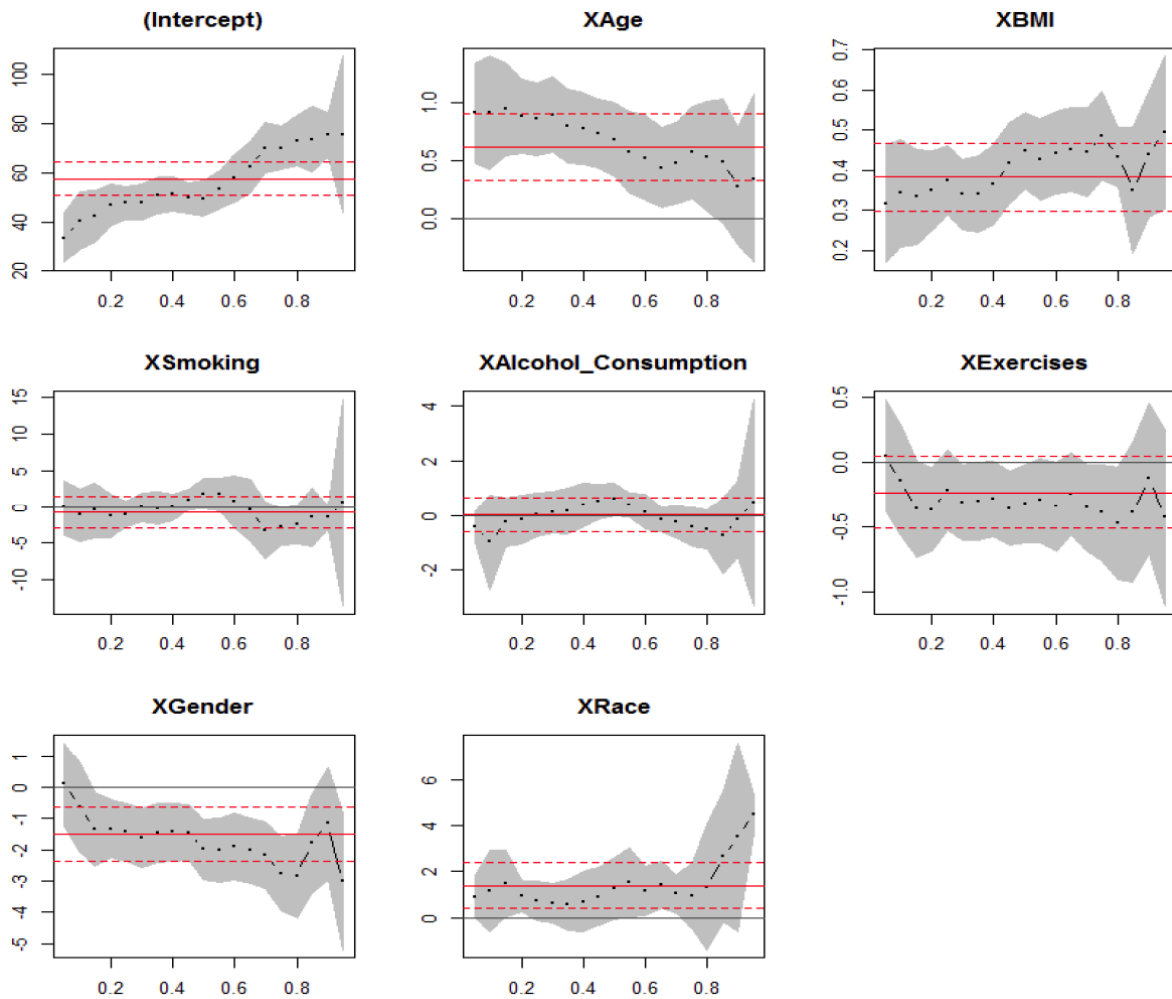$p^{**} < 0.05$ $p^{***} < 0.01$.

Figure 1: OLS and QR Estimates for Diastolic Blood Pressure.

presented positive statistically significant relations with SBP across all other quantiles except the 95% quantile.

In OLS regression analysis, participants` age (*p*<0.01), BMI (*p*<0.01) and gender (*p*<0.01) presented statistically significant relations with SBP. Just like Figure **1**, Figure **2** presents a clear graphic summary of the quantile regression results as each plot shows one of the independent variables in the quantile regression model. Also, Figure **2** results collaborate to those presented in Table **4**.

Impact of Age

As shown in Tables **3**, **4**, Figures **1** and **2**, age has positive statistically significant estimated OLS and quantile regression coefficients except for the 95th percentile, with both Diastolic Blood Pressure and Systolic Blood Pressure respectively. The magnitude of the association for DBP decreases from low to high quantiles whilst for SBP the magnitude of association fluctuates across the quantiles. This outcome shows

that the age effect on DBP is much larger at the lower end than upper end and central locations. Interpretation of the quantile regression estimates for a given quantile of SBP and DBP resembles that of ordinary least squares regression. For instance, in the case of age, the 10th quantile estimate of DBP is 0.91 (Table **3**). Meaning holding other 6 risk factors constant, a 1 year increase in age results in 0.91 mmHg increase in DBP.

Impact of BMI

As shown in Tables **3**, **4**, Figures **1** and **2**, BMI has positive statistically significant estimated OLS and conditional quantile functions with both the BP measures except the 95th quantile for SBP. The magnitude of the association for DBP increases from low to high quantiles whilst for SBP the magnitude of association fluctuates across the quantiles. This result implies that the BMI effect on DBP is much larger at the upper end than lower end and central locations.
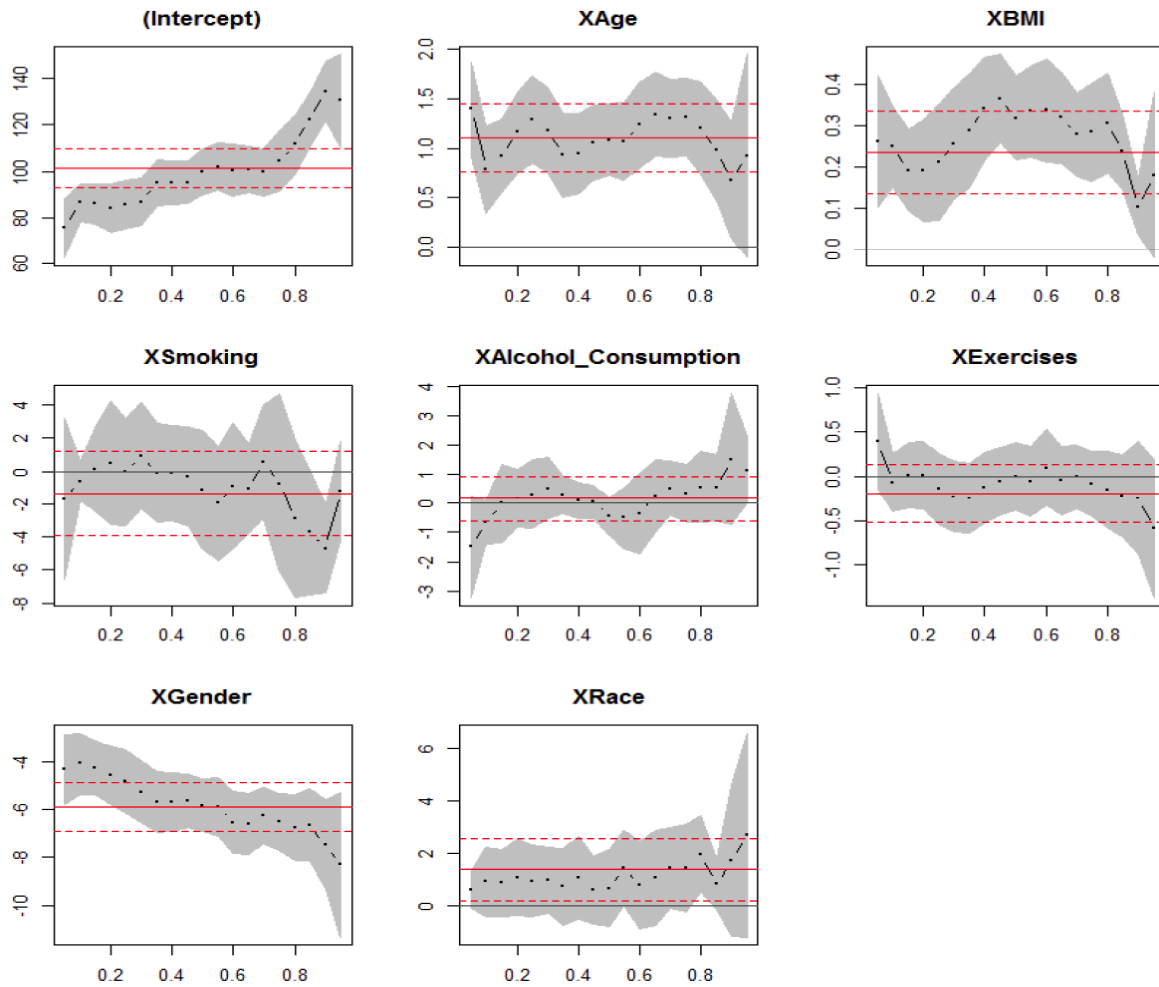
Figure 2: OLS and QR Estimates for Systolic Blood Pressure.

Impact of Smoking, Alcohol Consumption and Exercises

It is evident from Tables **3**, **4**, Figures **1** and **2** that smoking, alcohol consumption and exercises did not present significant relations with both DBP and SBP for all the estimated OLS and quantile regression coefficients at all significance levels respectively.

Impact of Gender

Gender is inversely related to both DBP and SBP except the 10th quantile for DBP. Table **3** and Figure **1** indicate that the DBP`s estimated coefficient on gender by the OLS method is equal to -1.49 whilst the estimated conditional quantile functions decreases from -0.65 in quantile 0.1 to -3.02 in quantile 0.95. Similarly, Table **4** and Figure **2** reveal that the SBP`s estimated coefficient on gender by the OLS method is equal to -5.92 whilst the estimated conditional quantile functions decreases from -4.08 in quantile 0.1 to -8.31 in quantile 0.95. These negative significant coefficients

entails that DBP or SBP is likely to increase more on males than females.

Impact of Race

Race has statistically significant estimated OLS and the 95th quantile with DBP (Table **3** and Figure **1**) whilst it is not statistically significant with the estimated OLS coefficient and estimated conditional quantile functions for SBP (Table **4** and Figure **2**).

DISCUSSION

As anticipated, risk factors (BMI, age, race, gender, alcohol consumption, smoking and exercises) differently affected study participants` DBP and SBP in each quantile.

BMI has positive statistically significant estimated OLS and conditional quantile functions with both the BP measures except the 95th quantile for SBP. The magnitude of the association for DBP increases from

low to high quantiles whilst for SBP the magnitude of association fluctuates across the quantiles. Age had also positive statistically significant estimated OLS and quantile regression coefficients except for the 95th percentile, with both DBP and SBP respectively. The magnitude of the association for DBP decreases from low to high quantiles whilst for SBP the magnitude of association fluctuates across the quantiles. These findings are almost similar to a study conducted by Kim *et al.* which showed that age and BMI had positive significant influence in the 5%, 25%, 50%, 75% and 95% quantiles [6]. Moreover, in a study by Cappuccio *et al.* [16], body mass index was found to be the most significant determinant of mean systolic and diastolic blood pressure.

Gender was found to be inversely related with both DBP and SBP except the 10th quantile for DBP. This result was confirmed by Shen *et al.* [15].

Smoking, alcohol consumption and exercises did not present any statistically significant relations with both DBP and SBP for all the estimated OLS and quantile regression coefficients at all significance levels respectively. These findings agree with findings reported by Soudarssanane *et al.* [17] in the urban adolescents study. No significant association was found between physical activity and blood pressure. Similary, Kim *et al.* [7] mentioned that the effects of physical activity were not significant in any quantile. No significant relations were found between blood pressure and smoking/alcohol consumption in adolescents [17].

Probably exercises did not present statistically significant association with both blood pressure measures because they are very few respondents who regularly participate in exercises (14.4%) for there to exist statistically significant difference. Likewise, smoking and alcohol consumption did not present statistically significant relations with both SBP and DBP because of small number of children with smoking (4.2%) and alcohol consumption (6.84%) habits to reveal out a significant change in both blood pressures.

CONCLUSION

In summary, BMI, age, gender and partially race were found to be predictors of blood pressure (both DBP and SBP) in the study participants using both ordinary least squares and quantile regression techniques. Considering that age, gender and race are risk factors of DBP and SBP that are beyond the control of human beings, there is nothing that one can do to manage these risk factors.

However, BMI can be managed by avoiding gaining excessive weight. Since most children in South Africa are not physically active as evidenced by 69.9% of the participants in this study not exercising at all, this could lead to most of them being overweight or obese. As a result, high levels of obesity may lead to high prevalence of DBP and SBP.

These study findings call for programs and policies meant to curb unhealthy lifestyles, promote physical activity and healthy nutrition among children. We recommend introduction of more sporting activities in schools and communities so that children can participate more in physical activities.

Furthermore, we recommend setting up of wellness centers at schools and communities where children can get tested regularly on BMI, blood pressure and other body conditions. Also, wellness centers may assist to inform children on risks of being overweight or obese and risks associated with abnormal blood pressure.

Last but not least, wellness centers could be used to run campaign services meant to promote healthy lifestyles among children. For instance, to abstain from smoking and drinking alcohol.

Introducing more health awareness talks on national television and radio channels may help children and adults in South Africa to lead healthy life styles and reduce mortality rates caused by abnormal blood pressure.

Finally, parents should also encourage their children to do outdoor activities and games instead of watching televisions and play computer games (sedentary activities) regularly during weekends and school holidays.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Christofaro DGD, Fernandes RA, Polito MD, Romanzini M, Ronque ERV, Gobbo LA, Oliveira AR de. A comparison between overweight cutoff points for detection of high blood pressure in adolescents. J Pediatria 2009; 85(4): 353-8. https://doi.org/10.2223/JPED.1911

[2]   Aounallah-Skhiri H, El Ati J, Traissac P, Ben Romdhane H, Eymard-Duvernay S, Delpeuch F, Achour N, Maire B. Blood pressure and associated factors in a north African adolescent population. A national cross-sectional study in Tunisia. BMC Public Health 2012; 98: 1-10.
https://doi.org/10.1186/1471-2458-12-98

[3]   Chiolero A, Madeleine G, Gabriel A, Burnier M, Paccaud F, Bovet P. Prevalence of elevated blood pressure and association with overweight in children of a rapidly developing country. Journal of Human Hypertension 2006; 21(2): 120-7.
https://doi.org/10.1038/sj.jhh.1002125

[4]   Feber J, Ahmed M. Hypertension in children: New trends and challenges. Clinical Science 2010; 119(4): 151-61.
https://doi.org/10.1042/CS20090544

[5]   Ayatollahi S, Vakili M, Behboodian J, Zare N. Reference Values for Blood Pressure of healthy schoolchildren in Shiraz (Southern Iran) using Quantile Regression. Iranian Cardiovascular Research Journal 2010; 4(2): 55-65.

[6]   Falkner B. Hypertension in children and adolescents: Epidemiology and natural history. Pediatric Nephrology 2009; 25(7): 1219-24.
https://doi.org/10.1007/s00467-009-1200-3

[7]   Kim HS, Park YH, Park HB, Kim SH. Estimation of effects of factors related to preschooler body mass index using Quantile regression model. Asian Nursing Research 2014; 8(4): 293-9.
https://doi.org/10.1016/j.anr.2014.07.005

[8]   Crispim PAA, Peixoto M do RG, Jardim PCBV. Risk factors associated with high blood pressure in Two- to Five-Year-Old children. Arquivos Brasileiros de Cardiologia 2014; 102(1): 39-46.

[9]   Peer N, Lombard C, Steyn K, Levitt N. High prevalence of metabolic syndrome in the black population of Cape Town: The cardiovascular risk in black South Africans (CRIBSA) study. European Journal of Preventive Cardiology 2014; 22(8): 1036-42.
https://doi.org/10.1177/2047487314549744

[10]   He X, Liang H. Quantile regression estimates for a class of linear and partially linear errors-in-variables models. Statistica Sinica 2000; 10: 129-40.

[11]   Beyerlein A, Toschke AM, von Kries R. Risk factors for childhood overweight: Shift of the mean body mass index and shift of the upper percentiles: Results from a cross-sectional study. International Journal of Obesity 2010; 34(4): 642-8.
https://doi.org/10.1038/ijo.2009.301

[12]   Leibbrandt M, Woolard I, de DeVilliers L. National Income Dynamic Study. Methodology: report on NIDS wave 1, technical paper no. 1. 2009. Available at: http://www.nids.uct.ac.za/publications/technical-papers/108-nids-technical-paper-no1/file (Accessed on 15 December 2016).

[13]   Koenker R, Hallock KF. Quantile regression: An introduction. 2000; 15(4). Available from: https://www.researchgate.net/publication/247312065_Quantile_Regression_An_Introduction. (Accessed on 22 November 2016).

[14]   IBM SPSS Missing Values 22. 1989; IBM Corporation. Available from: http://www.sussex.ac.uk/its/pdfs/SPSS_Missing_Values_22.pdf_(Accessed on 23 November 2016).

[15]   Shen X, Li K, Chen P, *et al.* Associations of blood pressure with common factors among lety-behind farmers in rural China: A cross-sectional study using quantile regression analysis. Medicine 2015; 94(2): e142.
https://doi.org/10.1097/MD.0000000000000142

[16]   Cappuccio FP, Micah FB, Emmett L, Kerry SM, Antwi S, Martin-Peprah R, Phillips RO, Plange-Rhule J, Eastwood JB. Prevalence, detection, management, and control of hypertension in Ashanti, West Africa. Hypertension 2004; 43(5): 1017-22.
https://doi.org/10.1161/01.HYP.0000126176.03319.d8

[17]   Soudarssanane MB, Karthigeyan M, Stephen S, *et al.* Key predictors of high blood pressure and hypertension among adolescents: A simple prescription for prevention. Indian J Community Med 2006; 31(3): 164-169.