Journal of Pharmaceutical Research International

33(15): 69-82, 2021; Article no.JPRI.66861
ISSN: 2456-9119
(Past name: British Journal of Pharmaceutical Research, Past ISSN: 2231-2919, NLM ID: 101631759)

Prevalence of Obesity and Trends of Body Mass Index in Azad Jammu and Kashmir

Kamran Abbas[1], Sajjad Ali Gill[2*], Rehan Ahmed Khan Sherwani[1], Tabinda Qazi[1], Naveed Feroze[1], Rana Adil[2] and Hira Shahid[2]

[1]Department of Statistics, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan.
College of Statistical and Actuarial Sciences, University of the Punjab, Lahore, Pakistan.
[2]Department of Sports Sciences and Physical Education, University of the Punjab, Lahore, Pakistan.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2021/v33i1531289
Editor(s):
(1) Dr. Mohamed Fathy, Assiut University, Egypt.
Reviewers:
(1) Carmine Finelli, Ospedale Cav. Raffaele Apicella, Azienda Sanitaria Locale Napoli 3 Sud, Italy.
(2) Shashi K. Agarwal, Agarwal Health Center, USA.
Complete Peer review History: http://www.sdiarticle4.com/review-history/66861

Original Research Article

Received 06 January 2021
Accepted 12 March 2021
Published 23 March 2021

ABSTRACT

In the present study an attempt has been made to figure out the remarkable factors that were responsible for higher body mass index (BMI) and obesity as well as to observe the existing trends of BMI in Azad Jammu and Kashmir. The factors were age, sex, diabetes, blood pressure, total cholesterol, triglycerides, blood glucose and random glucose. The statistical analysis with both classical and Bayesian methodologies was carried out for the investigation of significant risk factors. The results showed that with the increase in age the BMI also increases i.e., respondents of age group > 60 have the highest percentage of 82.76% for BMI. Moreover, it was found that except sex and blood pressure all other factors had significant association with BMI. Additionally, four factors namely, age, diabetes, total cholesterol and triglycerides were selected for the development of the parsimonious model of BMI based on generalized linear model, step wise regression and Bayesian model averaging.

Keywords: Body mass index; generalized linear model; logistic regression; bayesian model averaging; akaike information criterion; bayesian information criterion.

**Corresponding author: E-mail: sajjad.sspe@pu.edu.pk, drsajjadaligill@gmail.com;*

1. INTRODUCTION

Along with economic advancement and globalization the human beings have been facing many problems for few decades. The health related issues are believed to be the most concerning. Among these issues, obesity is considered as a major health problem in present days. Obesity rates have been increasing since last few decades. Excessive body fat is considered as a major issue for numerous diseases like diabetes, cancer and cardiovascular disorders. The prevalence of obesity and overweight have been reported to be continuously growing, not only in the Asian countries but in Western countries too. Obesity has been defined as a weight higher than 20% above what is considered to be normal according to standard age, height and weight tables. It can also be recognized by a formula known as the body mass index (BMI). The BMI may be defined as:

$$BMI=\frac{MASS(kg)}{HEIGHT(m^2)}$$

The BMI is considered an alternative for direct measures of a body fat. Generally, BMI is not only cheap but also easy to manage method of screening for weight categories that may later cause swear health problems. The BMI numeric is not same for male and female as any other biologic phenomenon. World Health Organization (WHO) in 2000 has represented different guidelines on BMI for universal population and South Asia. For universal population, WHO has fixed worth 25 for normal BMI, 25 to 30 for morbidly obese and above 30 for obesity. In contrast, WHO sets comparatively firm BMI levels for South Asian countries, which put value 23 for common BMI, 23 to 27.5 for overweight and above 27.5 for obese? In contrast, the cut off levels for normal BMI presented by Chinese (2002) are as 18.5 kg/ m^2 ≤ BMI < 24 kg/m^2 means less than 18.5 is underweight and higher than 24 is overweight and obesity categories. In the current study the cut off values taken for BMI are set by Chinese.

However, the risk of obesity can be minimized by timely modifying specific factors which are considered to be associated with higher BMI. As Pakistan is 6^{th} most populous country and according to recent research it has been reported that one in four adults is overweight/obese. The BMI is a good proxy to catch obesity on population level. Many articles on obesity have been published not long time ago. Ogden et al. [1] described national estimates of the frequent obesity and its drifts among United States children and adolescents availing the third National Health and Nutrition Examination Survey (NHANES-III) data. Based on the statistical analysis, they inferred that the universality of overweight among children in the United States is getting increased, especially in Mexican American and Non-Hispanic black youngsters. Moreover, Nandram et al. [2] exclusively scrutinized the NHANES-III data and applied a Bayesian method which used small area estimation methods (Rao, 2003) with a spline regression to estimate the percentiles at each age (2 to19 years). Qasim et at. (2014) conducted a study in Bhimber, AJK to purpose the total Cholesterol, Triglycerides and BMI of unmarried males and females. Moreover, Mtambo [3] proposed spatio temporal quantile interval regression models for childhood stunting, overweight, and obesity in Republic of Congo from 2005 to 2012 based on Demographic and Health Survey datasets. He conclude that mother's BMI had significant nonlinear effects on childhood overweight and obesity. Yu et al. [4] reviewed both classical and modern statistical methods for BMI analysis, highlighting that most of the classical methods were simple and easy to implement but ignore the complexity of data and structure, whereas modern methods did take complexity into consideration but could be difficult to implement. Terada et al. [5] analyzed data of 7560 patients who underwent coronary artery bypass grafting using BMI of 18.5 to 24.9 as a reference. They concluded that a greater attention and intervention to control the risks associated with infection and length of stay in patients with BMI ≥40.0 might improve patient care quality and efficiency.

Additionally, Martin et al. [6] evaluated the association between exercise and BMI in adults and concluded that there was no association between hours of exercise per week and BMI. Similarly, Flegal et al. [7] estimated the prevalence of adult obesity from 2009 to 2010 NHANES-III data and compared adult obesity and the distribution of BMI with data from 1999 to 2008. They found that trends in BMI were similar to obesity trends.

Moreover, Nandram et al. [2], Tesfaye et al. [8], Nishida et al. [9], Moens et al. [10], Cole et al. [11], Paeratakul et al. [12], Mitchel et al. [13] and Li et al. [14] investigated the relation between BMI and different risk factors. However, Acquah et al. [15], Yang et al. [16] and Raftery et al. [17] applied both classical and Bayesian methods to identify the significant risk factors of BMI. Based on the abovementioned details, the major objectives of the study are to:

- Determine which lifestyle related factors have the greater impact on obesity risk.
- Find the significant risk factors of BMI by using different statistical techniques.
- Develop the suitable model for BMI by using Generalized Linear models, Stepwise Logistic Regression, Bayesian Logistic Regression and Bayesian Model Averaging.

Including this introduction section, the rest of the article unfolds as follows. Section 2 contains the materials and methods whereas Section 3 comprises results and discussion. Section 4 designates the parsimonious model for BMI and Section 5 quantifies the summary and discussion.

2. MATERIALS AND METHODS

2.1 About the Data

The existent study conducted in Muzaffarabad, AJK. In which 300 respondents took part. The data was collected in collaboration with City Diagnostic Laboratory Muzaffarabad, Azad Jammu and Kashmir (AJK). Whereas, all the respondents have the ages between 20 to 81 year. About nine variables (factors) were observed against each respondent. The variables are age(AGE), sex(SEX), BMI, diabetes(DB), total cholesterol(TC), triglycerides(TG), blood pressure(BP), blood glucose(BG) and random glucose(RG). In parenthesis, the notation used for each variable during the statistical analysis has been given. Except age all the factors are binary such that they have only two categories i.e., 0 and 1. While '0' represents the normal value and 1 represents the abnormal. Whereas, age has been divided into three categories (20-40, 41-60, 61-81). BMI has been taken as the outcome variable and it has been coded as 0 and 1, representing 0 as less than 24kg/m^2 and 1 as higher than 24kg/m^2.

2.2 Statistical Methods

The statistical techniques applied in the current study are bivariate analysis (Chi-Square and Fisher Exact), Generalized Linear models, Stepwise Logistic Regression, Bayesian Logistic Regression and Bayesian Model Averaging. Different models and the corresponding odds ratios have also been computed to check the association between different risk factors and the outcome variable. The significance level selected was 0.05.

3. RESULTS AND DISCUSSION

The resulting distribution of BMI (whether 0 or 1) across all different factors is presented in Table 1. and Fig. 1. It is observed that respondents of age group > 60 have the highest percentage of 82.76% for BMI. While those belong to age group < 41 have lowest percentage of 20.29% for BMI. Additionally, it is also noticed that male respondents have higher percentage of 44.03% than of female respondents (41.57%) for higher BMI.

Moreover, it has also been observed diabetic and nondiabetic respondents have a significant difference between their percentages i.e., 66.67% and 40.29% respectively. Also, it is noted that respondents who have high level of TC have higher occurrence (70.24%) of higher BMI than those who have normal cholesterol level (31.94%). A matching trend has also been observed for TG, BG and RG. Furthermore, those who were suffering from high BP had higher percentage (42.86%) of BMI. all this discription is shown in Fig. 1. In which x-axis represents the risk factors whereas, y-axis denotes the percentages of the coressponding risk factors.

3.1 Bivariate Analysis

To find out the association between BMI and other risk factors, bivariate analysis was carried out. Goodness of fit statistics such as, Pearson Chi-square test, Fisher tests and odds ratio (OR) with 95% confidence interval (CI) were calculated and presented in Table 2. P-values (<0.05) of Chi-square and Fisher statistics depict that BMI has a significant relation with all variables except SEX and BP. While the results computed from OR (under 95% CI) show that BMI has a significant relation with all risk factors such that, all the values of odds ratios are greater than 1. Moreover, it can be seen clearly that the respondents belong to the age group > 61 has higher prevalence of obesity (OR=18.43) as compare to respondents having age 41 to 60 (OR=5.2). The diabetic respondents have the odds of being obese 19.5% greater than the non-diabetic respondents. Also, the subjects having abnormal level of TC have the higher odds of obesity than those having normal level of TC. A similar trend was observed for the TG, BG and RG levels. The graphs of ORs along with 95% confidence interval (CI) for eight independent variables are given in Fig. 2. and 3. respectively.

4. SELECTION OF THE PARSIMONIOUS MODEL

To develop the parsimonious model for BMI generalized linear model, step wise logistic regression and Bayesian models have been applied to identify the significant risk factors. In which the BMI is considered as a dichotomous dependent variable and rest of variables have been treated as independent. The results of these models have been summarized in the following subsections.

4.1 Generalized Linear Model

The results obtained by applying generalized linear model (GLM) are shown in Table 3. From the estimates, it can be seen clearly, that age, DB, TC, TG and BG are all influencing positively. Whereas, sex, BP and RG have a negative effect on BMI. Moreover, it is observed that except AGE and TG all the estimated coefficients have a non-significant relation with BMI as their p-values are greater than 0.05. Also, we observed that noticeable reduction in residual deviance (326.01 at 291 degrees of freedom (DF)). That is the deviance has been reduced by 83.4 points with the 8 DF. Usually, a model with lowest value of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) is selected. Here, we have only one model, so we cannot say much about them. Fig. 4. is depicting the distribution of the estimated coefficients of the predictor variables obtained through generalized linear model. It is noticed that along with intercept, the three predictors namely SEX, BP and RG are also showing the negative influence on BMI. On the other hand, rests of variables are having positive influence on the outcome variable.

4.2 Stepwise Logistic Regression

For the best fitted parsimonious model, we have also applied the stepwise regression model to BMI dataset and results are presented in Tables 4-5. Table 4 has shown all the steps of stepwise regression. Whereas, Table 4 and 5 are showing the final model analysis including odds ratios for each explanatory variable. From Table 4, it can be seen clearly that in stepwise regression model, variables are being dropped one by one on the basis of AIC value. That is, a variable with lowest AIC value has to be dropped from the selected regression equation. Moreover, the negative sign indicates the inclusion of variables in the model and vice versa. So, it is selecting four significant variables for BMI namely, AGE, SEX, TG and BG. All having negative sign in the final step. As the model 5 has the lowest AIC value so it has been selected as the final model. From Table 6, it can be seen that the age group > 61 has higher odds of obesity as its odds ratio is greater. Additionally, it is observed that the diabetic subjects have odds of being obese 2.09 times more than odds non-diabetic. Similarly, the odds of respondents having normal level of TG is 2.10 times less than the respondents having non normal level. A similar pattern has been observed for the risk factor TC.

4.3 Bayesian Logistic Regression Model

The Bayesian logistic regression model has been applied to the observed data by taking student t distribution as a prior distribution with 6 DF. To get efficient number of sample size based on Monte Carlo simulation technique, the results of posterior logistic regression model are presented in Table 7. There are 9 predictors including intercept in the model. It can be observed that the estimates of coefficients of Bayesian logistic regression model are approximately similar to the estimates derived from the generalized linear model. In Bayesian analysis we use credible intervals to assess the significance of the estimated coefficients of the predictors. Where the required credible interval is obtained by 2.5 and 97.5 percentiles values. The 2.5 and 97.5 percentiles are also given in the Table 6. Which shows that 95% of the sampled values fall between the credible interval values.

Moreover, the values of Monte Carlo Standard Error (MCSE) must be less than the 5% of the value of the standard deviation (SD) of the predictor's parameters [18]. In our study, it is very close to zero which means it is less than the 5% of the parameters' SD. Additionally, $\hat{R}$ =1, indicates the convergence of Markove Chain Monte Carlo (MCMC) simulation results. Additionally, the graphs in Fig. 5 are displaying distribution of the estimates of the parameters through the Bayesian logistic regression model. It is observed that these graphs are much alike to the distribution of estimates of the parameters measured through the generalized linear model. Moreover, the four variables (Age, DB, TC and TG) who have the positive influence on BMI are same as identified by the stepwise regression analysis. Moreover, Fig. 6. is representing the number of MCMC logit posterior sample obtained through the iteration technique. About 10000 times iteration have been run to acquire the convergence.

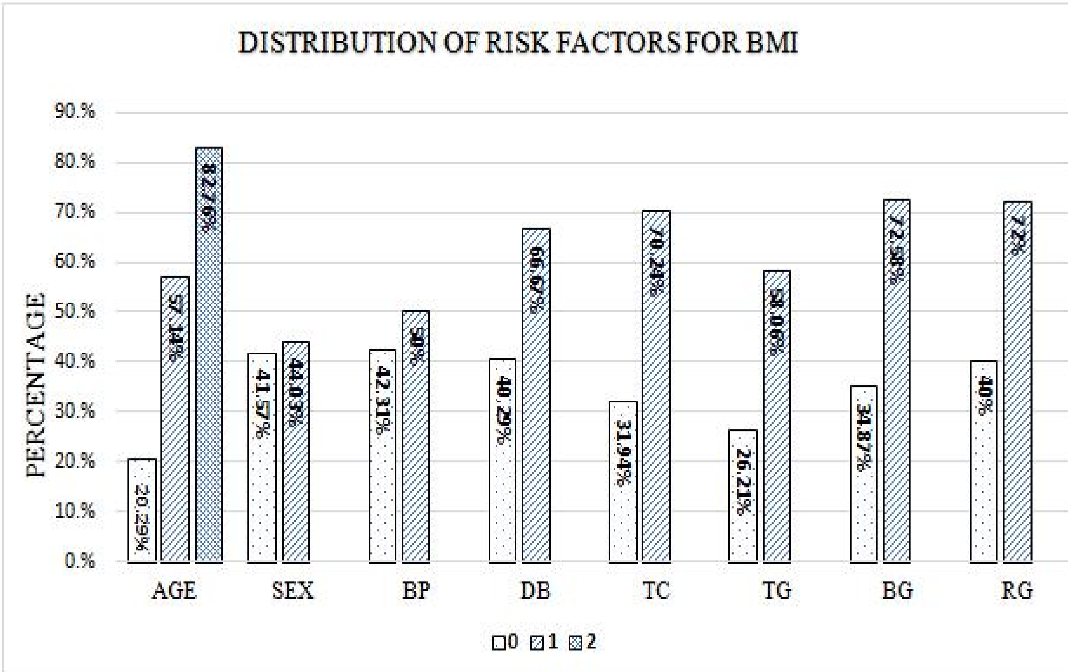

Fig. 1. Distribution of BMI risk factors

Table 1. Factors, categories, codes and distribution of different factors in the subjects of AJK for assessment of BMI

Factors (notation)	Categories	Codes	BMI		
			Yes (%)	No (%)	Total
Age (AGE)	20-40	0	28 (20.29)	110 (79.71)	138
	41-60	1	76 (57.14)	57 (42.86)	133
	61-81	2	24 (82.76)	5 (17.24)	29
Sex (SEX)	Female	0	69 (41.57)	97 (58.43)	166
	Male	1	59 (44.03)	75 (55.97)	134
Blood pressure (BP)	≤80-120	0	121(42.31)	165(57.69)	286
	<80-120	1	7(50.00)	7(50.00)	14
Diabetes (DB)	Non-diabetic	0	110 (40.29)	163 (59.71)	273
	Diabetic	1	18 (66.67) 9 (33.33)	9(33.33)	27
Total cholesterol (TC)	≤ 5.2mmol/l	0	69(31.94)	147(68.06)	216
	>5.2mmol/l	1	59(70.24)	25(29.76)	84
Triglycerides (TG)	≤1.7mmol/l	0	38(26.21)	107(73.79)	145
	>1.7mmol/l	1	90(58.06)	65(41.94)	155
Blood glucose (BG)	≤5.5mmol/l	0	83(34.87)	155(65.13)	238
	>5.5mmol/l	1	45(72.58)	17(27.42)	62
Random glucose (RG)	≤11.1mmol/l	0	110(40.00)	165(60.00)	275
	>11.1mmol/l	1	18(72.00)	7(28.00)	25

Table 2. Bivariate analysis of BMI and other risk factors

	Chi-square test			Fisher Test	Odds Ratio
Factors	χ^2	**df**	**P-value**	**P-value**	**OR[95% CI]**
Age	58.696	2	0.000*	0.000*	5.2 [2.69, 9.34] 18.43 [18.16, 67.44]
Sex	0.18	1	0.668	0.725	1.11 [0.68, 1.80]
DB	0.17	1	0.008*	0.013*	2.95 [1.21, 7.75]
TC	36.26	1	0.000*	0.000*	5.00 [2.81, 9.08]
TG	31.08	1	0.000*	0.000*	3.88 [2.33, 6.56]
BP	0.320	1	0.570	0.591	1.36 [0.40, 4.68]
BG	28.59	1	0.000*	0.000*	4.92 [2.57, 9.77]
RG	9.59	1	0.002*	0.003*	3.84 [1.47, 11.25]

Note: The asterisk (*) indicates the significant p-value (<0.05)

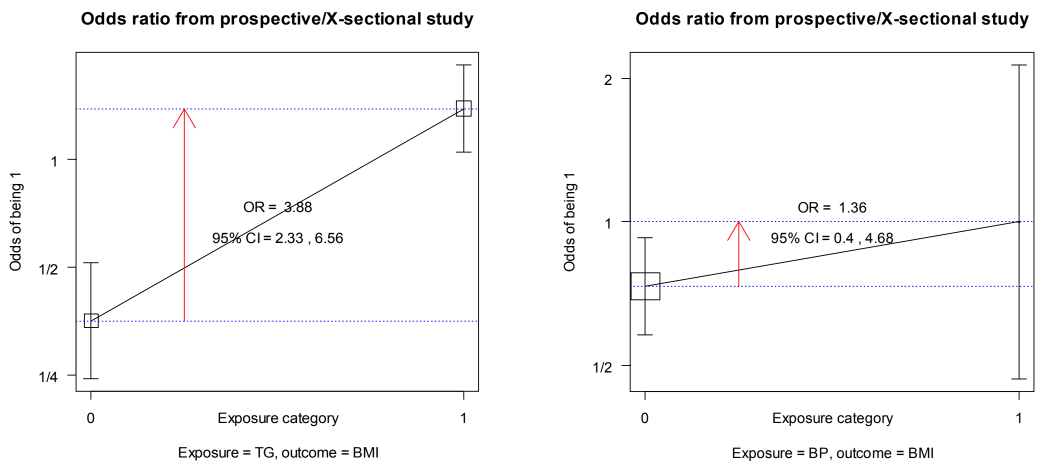

Fig. 2. Plots of OR of AGE, SEX, DB and TC with 95% confidence intervals

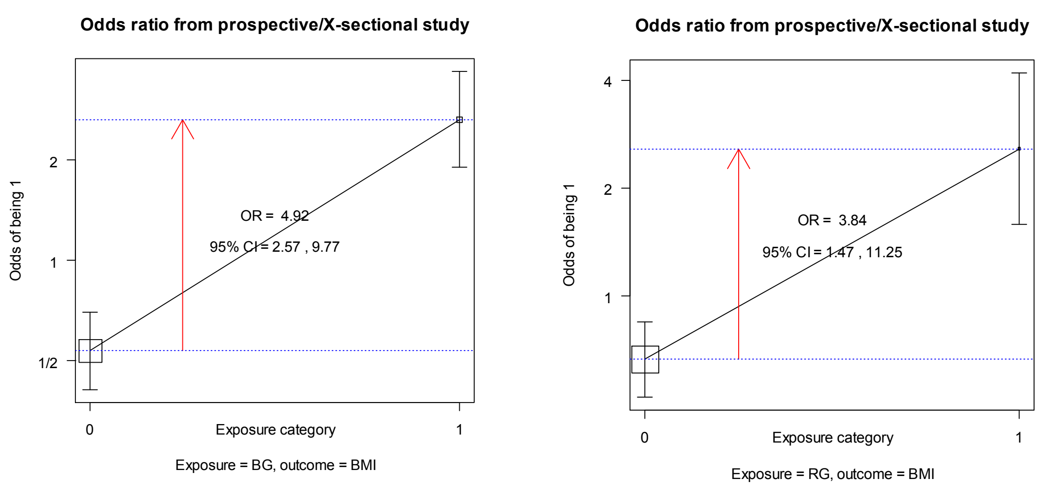

Fig. 3. Plots of OR of TG, BP, BG and RG with 95% confidence intervals

Table 3. Generalized Linear Model of BMI

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-1.730	0.272	-6.364	1.97e-10 ***
as.factor(AGE)1	1.283	0.302	4.244	2.20e-05 ***
as.factor(AGE)2	2.256	0.571	3.946	7.96e-05 ***
SEX	-0.154	0.277	-0.556	0.5781
DB	0.735	0.610	1.166	0.2435
BP	-0.306	0.674	-0.454	0.6498
TC	0.552	0.350	1.575	0.1152
TG	0.723	0.296	2.444	0.0145 *
BG	0.427	0.418	1.019	0.3080
RG	-0.074	0.676	-0.110	0.9124
Null deviance= 409.41, DF=299			Deviance Residuals:	
Residual deviance=326.01, DF=291			Min. 1Q Median 3Q Max.	
AIC=346.01 BIC=383.0519			-2.27 -0.791 -0.531 0.897 2.078	

*Note: Std. Error = Standard Error; Min=Minimum; Max = Maximum; *** shows p-value <0.0005*

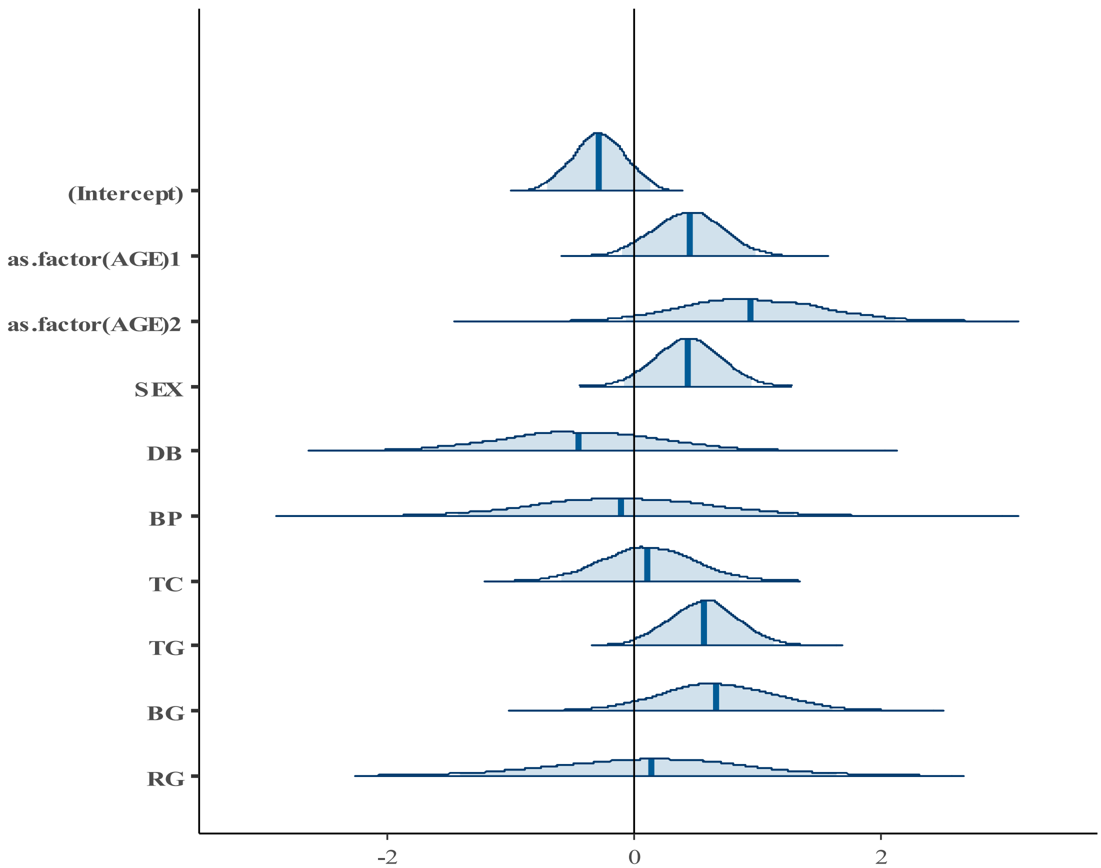

Fig. 4. Distributions of Coefficients of GLM model

Fig. 7 shows the posterior distribution of logistic regression model obtained through BMA. The spike at zero is showing the posterior probability given that the variable is not included in the model. On the other hand, the curve is exhibiting the model averaged posterior probability density of the estimated coefficients given that the variable is included in the model. As the density is scaled so the maximum height is approximately equal to the probability of the predictor being included in the model. Moreover, the figure depicts that out of eight only three predictors, namely AGE, TC and TG are being included in our final model. As the peaks of the DB and BG curves are too low that is approximately equal to zero, so we are not including DB and BG in our final model. It is also noticed that these variables (AGE, TC and TG) are similar to the variables obtained through the stepwise regression. But in stepwise regression model, DB is also included.

Table 4. Results of applying stepwise logistic regression

First step **AIC=346.01**

BMI ~ as.factor(AGE) + SEX + DB + BP + TC + TG + BG + RG

	Df	Deviance	AIC
- RG	1	326.03	344.03
- BP	1	326.22	344.22
- SEX	1	326.32	344.32
- BG	1	327.07	345.07
- DB	1	327.38	345.38
<none>		326.01	346.01
- TC	1	328.48	346.48
- TG	1	331.97	349.97
- as.factor(AGE)	2	352.95	368.95

Second step **AIC=344.03**

BMI ~ as.factor(AGE) + SEX + DB + BP + TC + TG + BG

	Df	Deviance	AIC
- BP	1	326.23	342.23
- SEX	1	326.33	342.33
- BG	1	327.16	343.16
- DB	1	327.64	343.64
<none>		326.03	344.03
- TC	1	328.48	344.48
+ RG	1	326.01	346.01
- TG	1	332.10	348.10
- as.factor(AGE)	2	352.95	366.95

Third step **AIC=342.23**

BMI ~ as.factor(AGE) + SEX + DB + TC + TG + BG

	Df	Deviance	AIC
- SEX	1	326.60	340.60
- BG	1	327.39	341.39
- DB	1	327.64	341.64
<none>		326.23	342.23
- TC	1	328.76	342.76
+ BP	1	326.03	344.03
+ RG	1	326.22	344.22
- TG	1	332.22	346.22
- as.factor(AGE)	2	353.04	365.04

Fourth step **AIC=340.6**

BMI ~ as.factor(AGE) + DB + TC + TG + BG

	Df	Deviance	AIC
- BG	1	327.79	339.79
- DB	1	327.94	339.94
<none>		326.60	340.60
- TC	1	329.20	341.20
+ SEX	1	326.23	342.23
+ BP	1	326.33	342.33
+ RG	1	326.59	342.59
- TG	1	332.45	344.45
- as.factor(AGE)	2	353.04	363.04

Last step **AIC=339.79**

BMI ~ as.factor(AGE) + DB + TC + TG

	Df	Deviance	AIC
<none>		327.79	339.79
- DB	1	330.16	340.16
+ BG	1	326.60	340.60
+ SEX	1	327.39	341.39
+ BP	1	327.49	341.49
+ RG	1	327.66	341.66
- TC	1	331.77	341.77
- TG	1	334.31	344.31
- as.factor(AGE)	2	358.09	366.09

Table 5. BMI Model using Stepwise Regression

Coefficients	Estimate	Std. Error	z-value	Pr(>|z|)
(Intercept)	-1.8003	0.2546	-7.072	1.53e-12 ***
as.factor(AGE)1	1.3189	0.2932	4.499	6.83e-06 ***
as.factor(AGE)2	2.3134	0.5646	4.098	4.17e-05 ***
DB	0.7399	0.4875	1.518	0.1291
TC	0.6668	0.3338	1.998	0.0457 *
TG	0.7438	0.2910	2.556	0.0106 *
Null deviance= 409.41, DF=299;		Residual deviance=327.79, DF=294		
AIC=339.79 ; BIC=362.01				

Table 6. Results of crude of the significant risk factors of BMI

Coefficients	Crude OR	95% CI
as.factor(AGE)1	3.74	[2.15, 6.64]
as.factor(AGE)2	10.12	[10.12, 30.57]
DB	2.09	[0.81, 5.49]
TC	1.95	[1.01, 3.75]
TG	2.10	[1.19, 3.72]

Table 7. Bayesian logistic regression model for BMI

Coefficients	Mean	SD	2.5%	97.5%	MCSE	Rˆ	n Effective
(Intercept)	-1.752	0.274	-2.308	-1.222	0.004	1.000	4316
as.factor(AGE)1	1.300	0.303	0.707	1.887	0.004	1.000	5485
as.factor(AGE)2	2.344	0.585	1.279	3.515	0.008	1.000	5418
SEX	-0.157	0.280	-0.706	0.384	0.004	0.999	4625
DB	0.754	0.657	-0.544	2.039	0.009	0.999	5215
BP	-0.339	0.713	-1.741	1.062	0.009	1.000	5780
TC	0.574	0.358	-0.125	1.269	0.005	0.999	5012
TG	0.735	0.297	0.164	1.320	0.004	1.000	4977
BG	0.454	0.433	-0.387	1.285	0.006	1.000	4873
RG	-0.067	0.714	-1.456	1.382	0.010	1.000	4691
mean_PPD	0.427	0.034	0.360	0.493	0.001	1.000	4306
log-posterior	-178.153	2.333	-183.692	-174.627	0.054	1.002	1841

Table 8. Bayesian model averaging

Coefficients	EV	SD	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	1.7e+00	0.271	-1.78	-1.76	-1.77	-1.48	-1.82
Age							
1	1.4e+00	0.297	1.48	1.32	1.33	1.39	1.47
2	2.5e+00	0.566	2.59	2.34	2.37	2.47	2.54
Sex	-3.7e-03	0.047	--	--	--	--	--
DB	9.2e-02	0.300	--	--	--	--	0.77
BP	-8.5e-05	0.082	--	--	--	--	--
TC	2.8e-01	0.428	--	0.69	--	0.99	--
TG	7.6e-01	0.418	0.97	0.76	0.86	--	0.94
BG	1.4e-01	0.329	--	--	0.70	--	--
RG	5.5e-02	0.239	--		--	--	--
BIC			-1353.8	-1352.4	-1352.1	-1351.2	-1350.84
post prob			0.312	0.157	0.136	0.086	0.070

Note: EV and SD mean the estimated variance and the standard deviation

4.4 Bayesian Model Averaging (BMA)

To develop the best fitted parsimonious model with significant predictors via Bayesian approach, BMA for logistic regression model has been applied to the observed BMI dataset. Unlike stepwise regression, BMA does not take different steps to exclude non-significant factors but takes different models with different predictors and gives the final models with significant predictors for a dichotomous outcome variable. In the present study, BMI has been selected as an outcome variable while all other variables are taken as predictors. The results of BMA for posterior logistic regression models are presented in Table 8. and Fig. 7. About 14 models have been selected and 5 best models are presented.

4.5 Parsimonious Models Obtained by Different Criterion

We have applied different techniques to develop the parsimonious models for BMI. The proposed models are presented in Table 9. From Table 9, it has been seen that the estimated intercept coefficients of all three models are negative. Whereas, the estimated values of slope coefficients are positive. Moreover, it has been observed all three methods have identified AGE and TG as the significant risk factors. While TC is identified through stepwise regression and BMA. While DB is only recognized through Stepwise Regression technique.

4.6 Comparison of Odds Ratios

The ORs calculated through different methodologies are presented in Table 10. It can be observed that ORs of only four significant factors are presented for the stepwise regression model, whereas, for all other techniques the ORs of those factors are also displayed which are statistically insignificant. Moreover, bringing light to the overall results of Table 10. it can be clearly seen that there is negligible difference between the ORs of bivariate analysis, Generalized Linear model and Bayesian Logistic regression analysis. Whereas, the ORs of Stepwise regression are different from other methods.

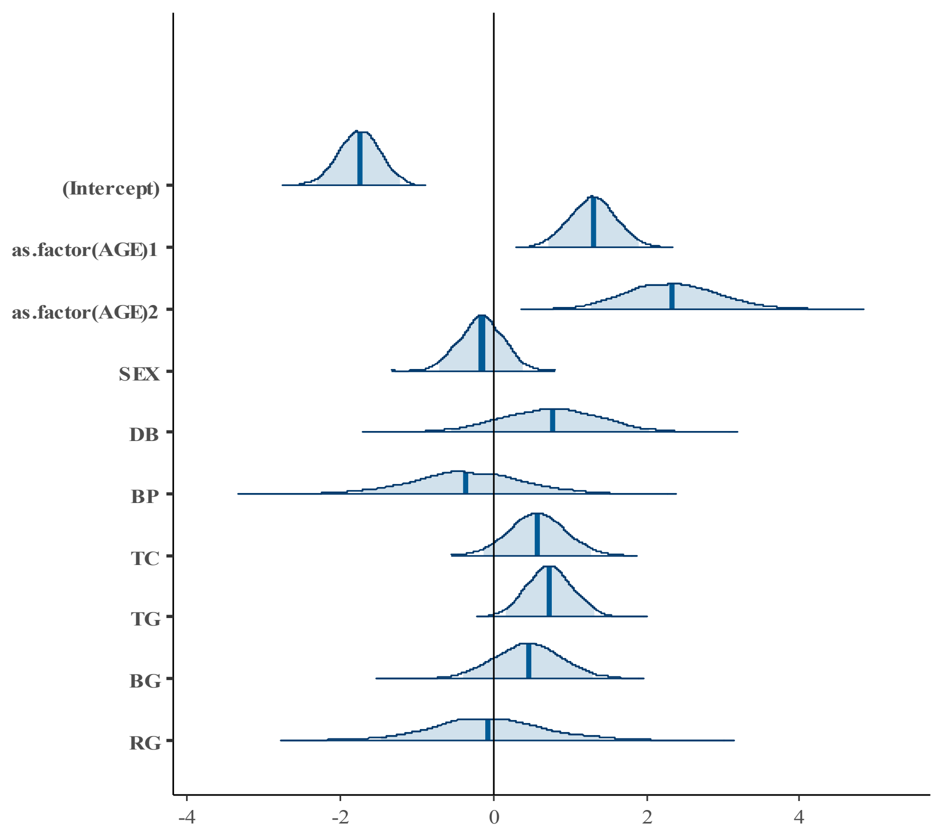

Fig. 5. Distribution of the Coefficients estimates via Bayesian Logistic Regression Model

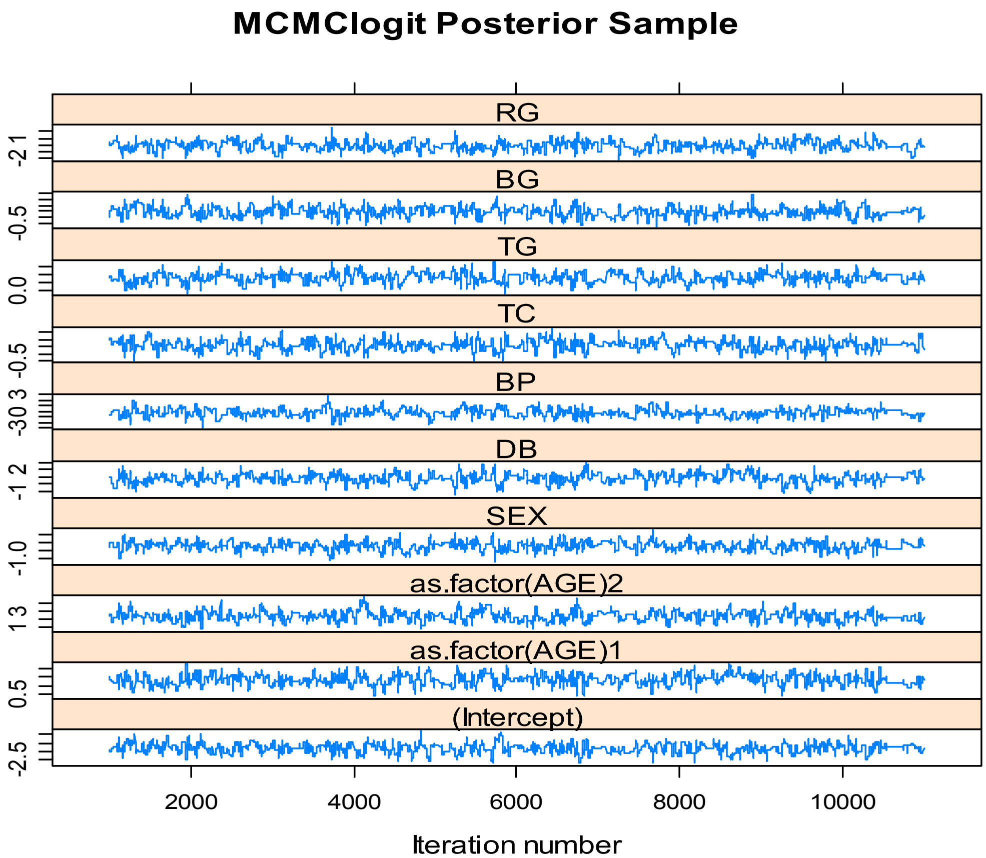

Fig. 6. Display of MCMC Logit Posterior Sample

Table 9. Parsimonious models for BMI

Methods	Models
	Generalized linear Model
	BMI~βo+β_1(AGE1)+ β_2(AGE2)+ β_3(TG)
	βo=-1.731 β_1 =1.283 β_2=2.256 β_3=0.724
	Stepwise Regression
	BMI~βo+β1(AGE1)+ β_2(AGE2)+ β_3(DB)+ β_4(TC)+ β_5(TG)
	βo=-1.800 β_1=1.318 β_2= 2.313 β_3=0.739 β_4=0.667 β_5= 0.743
	Bayesian Model Averaging
	BMI~βo+β1(AGE1)+ β_2(AGE2)+ β_3(TC)+ β_4(TG)
	βo= -1.76 β_1=1.32 β_2=2.34 β_3=0.69 β_4=0.76

Table 10. Comparison of ORs obtained through different methods

Factors	Bivariate analysis	Glm	Stepwise	Bayesian logistic
Age1	5.2	5.238	3.74	5.122
Age 2	18.43	18.857	10.12	17.418
Sex	1.11	1.11	--	1.107
DB	2.95	2.96	2.09	2.952
BP	1.36	1.36	--	1.336
TC	5.00	5.03	1.95	5.032
TG	3.88	3.9	2.10	3.910
BG	4.92	4.94	--	4.857
RG	3.84	3.86	--	3.825

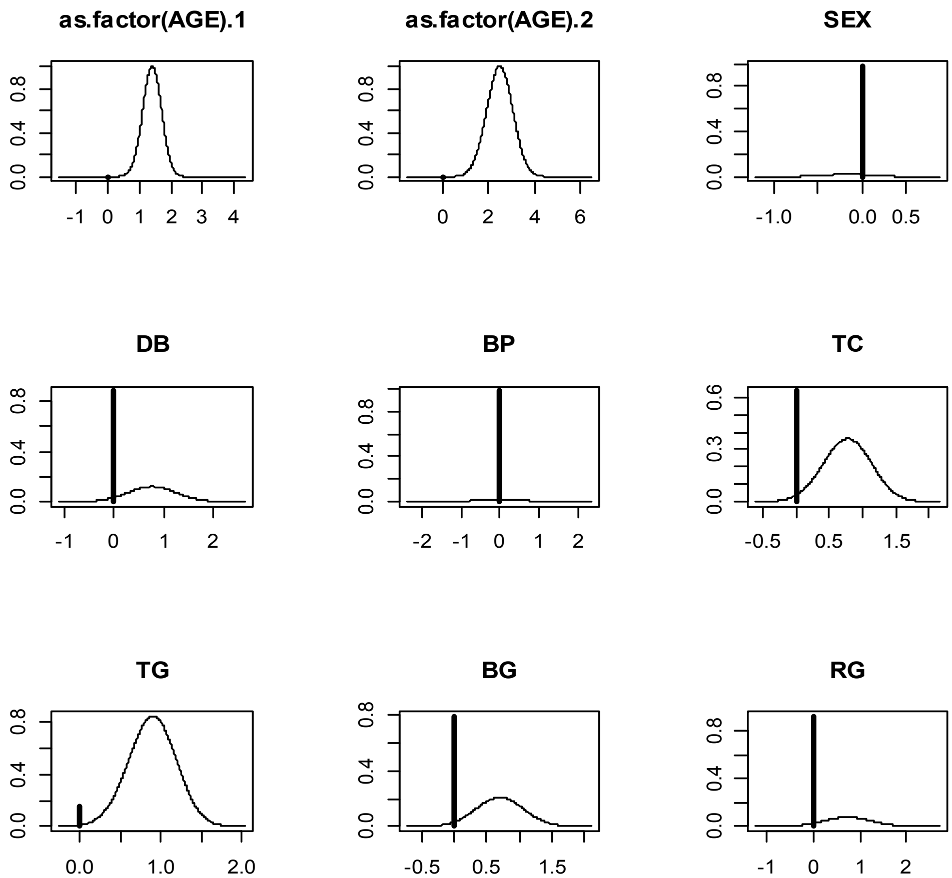

Fig. 7. BMA posterior distributions of the parameters of the logistic regression model

4. CONCLUSION

In modern era, obesity has become one of the basic health problems. The basic purpose to conduct this study is to explore the risk factors related to higher BMI and obesity. The secondary data used in the present study has been collected from City Diagnostic Lab, Muzaffarabad, AJK . There are nine variables including the outcome variable. The variables are age(AGE), sex(SEX), body mass index(BMI), diabetes(DB), total cholesterol(TC), triglycerides(TG), blood pressure(BP), blood glucose(BG) and random glucose(RG). In parenthesis, the notation used for each variable during the statistical analysis has been given. Initially, all the variables are coded then different statistical methods have been applied to find out different risk factors for the outcome variable (BMI) and their association with BMI.

Bivariate analysis has been carried out at first and the relative p-values for Chi-square and Fisher Exact tests have been calculated. It has been observed that sex and BP are found to be insignificant as their p-values are exceeding 0.05. Moreover, the ORs have also been computed. All the variables showed the positive association with BMI as the values of ORs were greater than 1. The Generalized Linear model analysis carried out by taking logit as a link function. The p-values calculated for all predictor variables showed that only age and TG is having a significant relation with BMI. The similar results have been obtained through Bayesian Logistic regression analysis and it has been noticed that Bayesian Analysis provided more stable results as the values of standard error for estimated coefficients were lower than that calculated via Generalized Linear Model.

Stepwise regression approach has also been used to find out the parsimonious model. A parsimonious model is one that neither underfits nor overfits i.e., having a reasonable set of independent variables. Four predictor variables were found to be significant for the model development via Stepwise regression. While the BMA gave three significant variables. Whereas, three variables namely, AGE, TG and TC are same identified through both classical Stepwise regression and BMA approach. While the fourth variable identified by Stepwise regression is DB. From a class of models, the selection of the parsimonious model depends upon some criterions like AIC, BIC or DIC (in Bayesian). The model with lowest AIC or BIC values is said to be parsimonious model. In current statistical analysis, the model with lowest BIC value is one that obtained through Stepwise regression analysis. So, the final selected model has four significant risk factors namely, AGE, DB, TC and TG.

OR is used to find out the association between two variables. It can be computed through different statistical methods like Chi-square, Regression analysis etc. All the factors have ORs greater than unity showing positive association between risk factors and the outcome variable (BMI). Whereas, it can be clearly seen that there exists a negligible difference among the ORs calculated through bivariate analysis, Generalized Linear Model analysis and the Bayesian Logistic regression analysis. These three methods have same results of odds.

CONSENT

As per international standard or university standard, respondents' written consent has been collected and preserved by the author(s).

ETHICAL APPROVAL

As per international standard or university standard written ethical approval has been collected and preserved by the author(s).

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFRENCES

1. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of obesity and trends in body mass index among United States children and adolescents, 1999-2010. Journal of the American Medical Association. 2012; 307(5):483-490.
2. Nandram B, Choi JW. A Bayesian analysis of body mass index data from small domains under nonignorable nonresponse and selection. Journal of the American Statistical Association. 2010; 105(489):120-135.
3. Mtambo OP. Analysis of childhood malnutrition in Republic of Congo. International Journal of Statistics and Applications. 2018;8(4):153-166.
4. Yu K, Alhamzawi R, Becker F, Lord J. Statistical methods for body mass index: a selective review of the literature. Statistical Methods in Medical Research, 2016;27(3): 789-811.
5. Terada T, Johnson JA, Norris C, Padwal R, Qiu W, Sharma AM, et al. Severe obesity is associated with increased risk of early complications and extended length of stay following coronary artery bypass grafting surgery. Journal of the American Heart Association. 2016;5(6):1-10.
6. Martin ISM, Vilar EG, Barato VP. Exercise and Body Mass Index: Are those two parameters related in adults? Journal of Negative and No Positive Results. 2016; 1(1):36-41.
7. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among United States adults, 1999 - 2010. Journal of the American Medical Association. 2012;307(5):491-497.
8. Tesfaye F, Nawi NG, Van Minh H, Byass P, Berhane Y, Bonita R, et al. Association between body mass index and blood pressure across three populations in Africa and Asia. Journal of Human Hyper tension. 2007;21(1):28-37.
9. Nishida N, Tanaka M, Hayashi N, Nagata H, Takeshita T, Nakayama K, et al. Determination of smoking and obesity as periodontitis risks using the classification and regression tree method. Journal of Periodontology. 2005;76(6):923-928.
10. Moens G, Gaal LV, Muls E, Viaene B, Jacques P. Body mass index and health among the working population: epidemiologic data from Belgium. The European Journal of Public Health. 1999; 9(2):119-123.
11. Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the United Kingdom, 1990. Archives of Disease in Childhood. 1995;73(1):25-29.
12. Paeratakul S, Lovejoy JC, Ryan DH, Bray GA. The relation of gender, race and

socioeconomic status to obesity and obesity comorbidities in a sample of US adults. International Journal of Obesity. 2002;26(9):1205-1210.
13. Mitchell UH, Johnson AW, Adamson B. Relationship between functional movement screen scores, core strength, posture, and body mass index in school children in Moldova. The Journal of Strength & Conditioning Research. 2015;*29*(5):1172-1179.
14. Li L, Gower BA, Shelton RC, Wu X. Gender specific relationship between obesity and major depression. Frontiers in endocrinology. 2017;292(8):1-6.
15. Acquah HD. Bayesian logistic regression modelling via Markov Chain Monte Carlo algorithm. Journal of Social and Development Sciences. 2013;4(4):193-197.
16. Yang F, Niu XF, Lin J. Bayesian Mixed-effects Polychotomous Response Model with Application to Diverse Population Collaboration (DPC) Data. Journal of Biometrics and Biostatistics, 2017;8(2):346-355.
17. Raftery AE, Painter IS, Volinsky CT. BMA: An R package for Bayesian model averaging. The Newsletter of the R Project. 2005;5(2):2-6.
18. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression. 2013;398. John Wiley and Sons

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/66861