Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 201 (2025) 96-113 EISSN 2392-2192

A Multiple Regression Model to Predict Mortality Among Patients Affected During Covid-19 in Some Developing Countries By

Mohammad Shakil[1], Rakhshinda Jabeen[2], J. N. Singh[3], Aneeqa Khadim[4], Musaddiq Sirajo[5], Mohammad Ahsanullah[6], and B. M. G. Kibria[7]

[1]Department of Mathematics, Miami Dade College, Hialeah, FL, USA

[2]Department of Medicine, Dow University of Health Sciences, Karachi, Pakistan

[3]Department of Mathematics & Computer Sciences, Barry University, Miami Shores, FL, USA

[4]Department of Statistics, Mirpur University of Science & Technology, Mirpur, Pakistan

[5]Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

[6]Department of Management Sciences, Emeritus Professor, Rider University, NJ, USA

[7]Department of Mathematics & Statistics, Florida International University, Miami, FL, USA

[1-7]E-Mail Address: mshakil@mdc.edu, rakhshinda.jabeen@duhs.edu.pk, jsingh@barry.edu, Aneeqa89@gmail.com, musaddiqsirajo@gmail.com, ahsan@rider.edu, kibriag@fiu.edu

ABSTRACT

Multiple regression analysis is one of the widely used multivariate statistical techniques employed in medical, health, and other applied research and studies for determining the correlation between a response variable and some combination of two or more predictor variables. The objective of our investigation in this paper is to evaluate the association of some risk factors with mortality (death) among patients during Covid-19. A multiple linear regression model is developed to analyze the impact of the body mass index, death rate, life expectancy and population on mortality (death) among patients affected during Covid-19 based on a sample data of some developing countries. It is observed that in the presence of the life expectancy and population, the body mass index and the death rate were good predictors of the mortality (death).

(Received 10 January 2025; Accepted 20 February 2025; Date of Publication 6 March 2025)

2020 Mathematical Sciences Classification: 65F359, 15A12, 15A04, 62J05.

Keywords and phrases: Body mass index, Covid-19, Death rate, Multiple regression, Obesity.

1. INTRODUCTION

Multiple linear regression is one of the most widely used statistical techniques in medical, biological, health, nursing, and other applied sciences, and have been studied by many authors and researchers. For example, the interested readers are referred to Neter et al. [15], Draper and Smith [16], Tamhane and Dunlop [17], Mendenhall and Sincich [18], Chatterjee and Hadi [19], Montgomery et al. [20], Suárez et al. [21], Cleophas and Zwinderman [22], Guzman and Kibria [23], and Johnson and Wichern [24], among others, for details on multiple regression analysis and its applications. For recent developments on linear and non-linear regression models, we refer to Kamel and Abonazel [25], and Kibria [26]. Multiple linear regression is defined as a multivariate technique for determining the correlation between a response variable Y and some combination of two or more predictor variables, X. For example, it can be used to analyze data from causal-comparative, correlational, or experimental research. It can handle interval, ordinal, or categorical data. In addition, multiple regression provides estimates both of the magnitude and statistical significance of relationships between variables. The objective of our investigation in this paper is to develop a multiple linear regression model to analyze the impact of some risk factors, such as the body mass index, death rate, life expectancy, and population on mortality (death) among patients during Covid-19.

Acute respiratory syndrome, after emergence in Wuhan, China, in late December 2019, (Zhou, Yu, Du, Fan et.al. [1]), is given the name of COVID-19 in February 2020, (Organization WH [2]). It was declared as pandemic by the World Health Organization in March 2020, (World Health Organization [3]). Since the beginning of the illness, it was considered as an illness with high mortality and was associated with various risk factors. While discussing different morbidities, common diseases are linked with it including Diabetes Mellitus, Cardiac issues, various cancers, and chronic kidney disease. Apart from chronic diseases, obesity was also taken as one of the major risk factors leading to poor clinical outcome and mortality. Obesity is defined as excessive fat accumulation that is associated with a risk to health. According to WHO, standard body mass index (BMI) over 25 kg/m^2 is overweight, and over 30 kg/m^2 is the start of obesity. The body mass index (BMI) is defined as a measure of body size and for weight-related health risk. It combines a person’s weight with their height, defined as follows:

Definition 1.1: The body mass index (BMI) is defined as a measure of body size and for weight-related health risk. It combines a person's weight with their height. It can be calculated using the following formulas:

(1.1) $BMI = Weight(kg)/[height(m)]2,$

(1.2) $BMI = Weight(lb)/[height(in)]2 \times 703.$

Thus, the results of a *BMI* measurement can give an idea about whether a person's weight is correct with respect to their height. Moreover, the *BMI* of a person can indicate whether they are underweight or if they have a healthy weight, or excess weight, or obesity, (**Table 1.1**, "Body mass index", [32]).

Table 1.1. Categorization of BMI

(Source: https://medlineplus.gov/ency/article/007196.htm: "Body mass index", [32])

BMI kg/m ^{2}	**CATEGORY OF WEIGHT**
Below 18.5	Underweight
18.5 to 24.9	Healthy
25.0 to 29.9	Overweight
30.0 to 39.9	Obese
Over 40	Extreme or high-risk obesity

Furthermore, as reported by https://www.weightwatchers.com/us/science-center/bmi-calculator, there appears to be an exponential relationship between BMI and mortality rate which is illustrated in the following

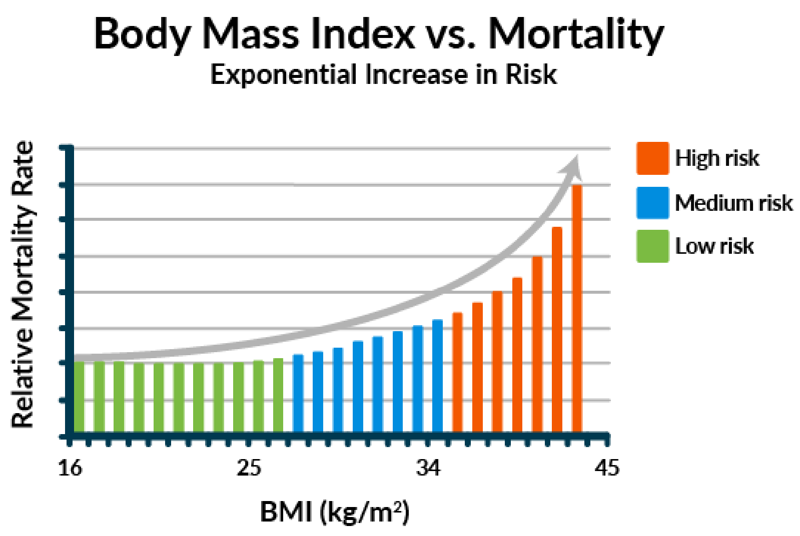

Figure 1.1
(Source: https://www.weightwatchers.com/us/science-center/bmi-calculator)

It is further categorized as obesity class 1, 2 and 3, (**Table 1.2** Obesity [4]).

Table 1.2
(https://medlineplus.gov/ency/article/007196.htm)

CLASS	OBESITY
1	BMI of 30 to less than 35.
2	BMI of 35 to less than 40.
3	BMI of 40 or higher. Class 3 is considered "severe obesity".

With the beginning of COVID-19, obesity was found to be major risk factor for severe illness and high mortality, (Patel and Jernigan [9], Kompaniyes, Goodman, Belay et.al. [10]). There are various studies done in different parts of the world in which obesity was significantly correlated with poor outcomes. A study done in Saudi Arabia shows higher mortality of 10.4% when compared with overweight patients (3.8%), (Habib, Aislim, Aurbidi et.al. [11]. Another study was carried out in New York, which also endorsed the same result, (Palaiodimos, Kokkinidis, Li et.al. [12]). The focus of this analysis is to evaluate association of obesity with mortality among patients with COVID-19. Overall risk of development of COVID-19 among obese patients was 46% and mortality was 48% more than the rest of the population, (Popkin, Du, Green, Beck et.al. [13]).

Obesity is directly correlated with mortality among patients with severe disease, which diseases proven by various studies. The mortality was up to four-fold and higher was observed in many countries (Gabbrielli and Pugno [14]).

Since the BMI is an important risk factors affecting the cardiovascular health issues of human beings, the purpose of the present study is to contribute to the applications of the multiple linear regression of the impact of the BMI and other factors such as the death rate, life expectancy, and population on mortality based on a sample data of some countries affected during Covid-19. It appears from the literature that not much attention has been paid to this kind of studies in the multiple regression analysis of the cardiovascular health issues and other problems in human population. Motivated by these facts, in this paper, a multiple linear regression model is developed to analyze the impact of the BMI and other factors such as the death rate, life expectancy, and population on mortality based on a sample data of some countries affected during Covid-19. The use of multiple linear regression is illustrated in the prediction study of mortality in human population affected during Covid-19 based on their BMI and other indicators such as the death rate, life expectancy, and population.

The organization of this paper is as follows. In Section 2, the proposed multiple linear regression model and the problem and objective of this study are presented. Section 3 provides the data analysis, justification, and adequacy of the multiple regression model developed. In Section 4, the residual plots of the proposed regression model are analyzed, and the goodness of fit of the model based on these residual plots is examined. Some concluding remarks are given in Section 5**.**

2. MULTIPLE LINEAR REGRESSION MODEL

2.1. A Multiple Linear Regression Model based on a Number of Predictor

Consider following multiple linear regression model

$$Y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2.1}$$

where Y is an $n \times 1$ vector of response variable (observations), β is a $k \times 1$ vector of unknown regression coefficients, X is an $n \times k$ ($n > k$) observed matrix of the regression, and ε is an $n \times 1$ vector of of random errors, which is distributed as multivariate normal with mean 0 and covariance matrix $\sigma^2 I_n$, and I_n is an identity matrix of order n. The OLS estimator of β is obtained as $\hat{\beta} = (X'X)^{-1}X'y$, and covariance matrix of $\hat{\beta}$ is obtained as $\mathrm{Cov}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$.

2.2. Problem And Objective Of Study

In what follows, a multiple linear regression model is developed to study the impact of the mean BMI and other factors such as the death rate, life expectancy, and population on mortality (death) based on a sample data of 36 countries affected during Covid-19, belonging to developing countries, (see APPENDIX 1), among 181 number of countries analyses done before, (Gabbrielli and Pugno [14]).

In these parts of the world, people are generally less obese than western society, and overal,l it has been observed that there is much smaller number of deaths among them. By selecting population with low BMI and death rate during Covid-19, we believed that it would be helpful in further investigations and comparing our results with the impact of BMI and other factors on the death in advanced countries during COVID-19 or other epidemics. For some recent studies on the applications of regression analysis to the number of deaths during Covid-19, the interested reads are referred to Siegrist and Kibria [27], Ghosal et al. [28], Kibria and Urbistondo [29].

For our sample data of 36 countries affected during Covid-19, belonging to developing countries, (see APPENDIX 1), the descriptive statistics of the response variable and the predictor variables are given in **Table 2.1**.

Table 2.1. Descriptive Statistics for Response Variable and the Predictor Variables.

Variable	Sample Size	Min	Max	Mean	Median	Standard Deviation
Death (Y)	36	1	130	59.33333	62	36.93005
Mean BMI (X_1)	36	21.6	27.5	24.12222	24.05	1.566271
Death Rate (X_2)	36	0.00036	1.26317	0.0642431	0.010735	0.2108346
Life Expectancy (X_3)	36	53.28	83.73	66.63611	65.015	8.498655
Population (X_4)	36	79843	9.818686e+7	1.264357e+7	5.440919e+6	2.161394e+7

The percentage breakdowns of the response variable, the death (Y), and the respective four predictor variables, the mean BMI (X_1), death rate (X_2), life expectancy (X_3), and population (X_4), are provided in Figures 2.1 – 2.5.

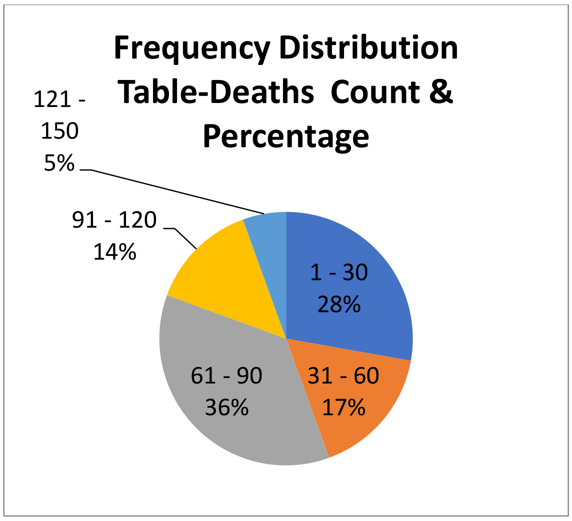

Figure 2.1

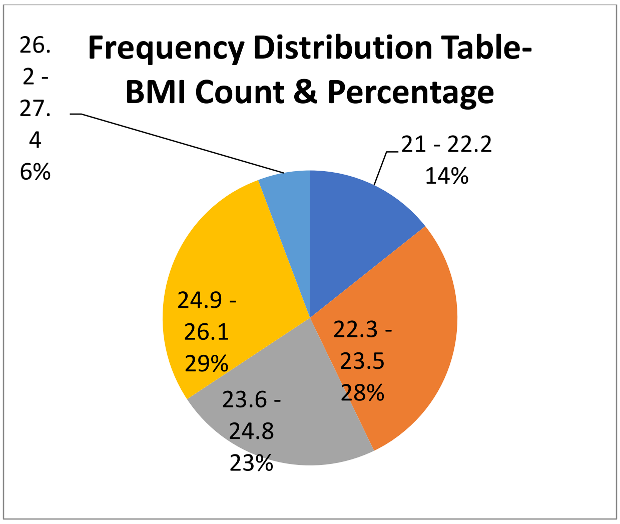

Figure 2.2

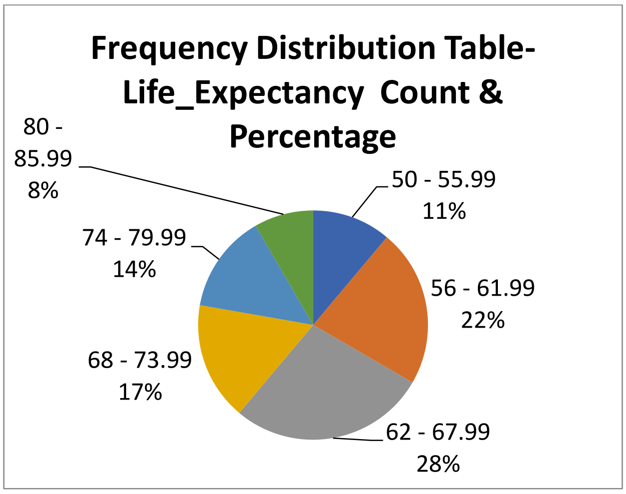

Figure 2.3

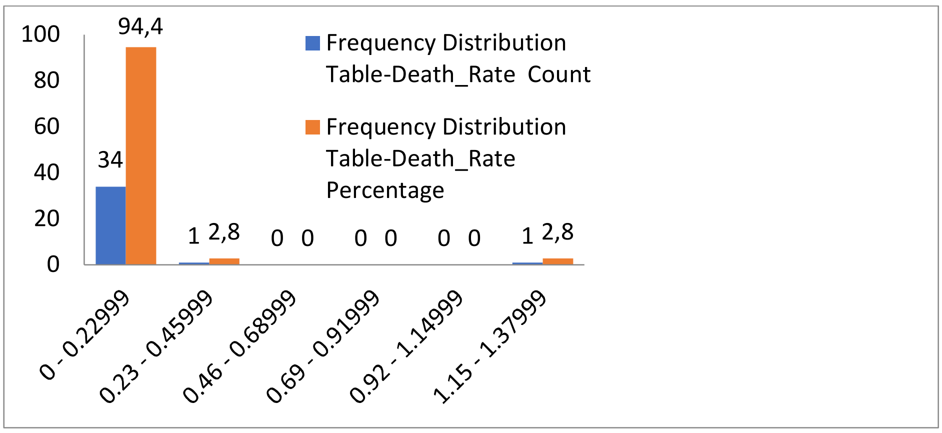

Figure 2.4

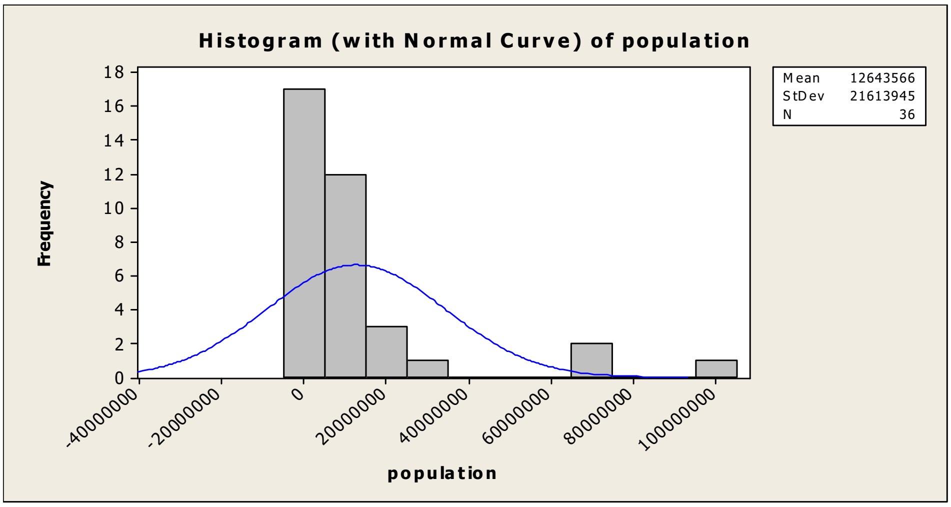

Figure 2.5

These figures are self-explanatory, as revealed by the respective pie charts and histograms, respectively. One can easily draw inferences from the basic data analysis as provided in these pie charts and histograms.

Using the Equation (2.1), the following four-predictor multiple linear regression first-order model (or the least squares prediction equation) was developed:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4, \tag{2.2}$$

where $\beta's$ denote the population regression coefficients, the response variable is the death $(\boldsymbol{Y})$, and the respective four predictors are the mean BMI $(\boldsymbol{X_1})$, death rate $(\boldsymbol{X_2})$, life expectancy $(\boldsymbol{X_3})$, and population $(\boldsymbol{X_4})$. The model in Equation (2.2) assumes that the relationship between the death $(\boldsymbol{Y})$ and each independent variable is linear, and the effects of each $\boldsymbol{X}$ on $\boldsymbol{Y}$ is independent of the other $\boldsymbol{X's}$ (that is, no interaction).

3. ANALYSIS OF DATA

The four-predictor multiple linear regression first-order model (2.2) was fitted to the Table A.1 in Appendix 1 using the Minitab Version 17.0 regression computer programs to determine the regression coefficients and analyze the data. The adequacy of the multiple linear regression model for predicting the death $(\boldsymbol{Y})$ based on the mean BMI $(\boldsymbol{X_1})$, death rate $(\boldsymbol{X_2})$, life expectancy $(\boldsymbol{X_3})$, and population $(\boldsymbol{X_4})$ was conducted using the F-test for significance of regression. For details on the analysis of the regression coefficients, etc., the interested readers are referred to Neter et al. [15], Draper and Smith [16], Tamhane and Dunlop [17], Mendenhall and Sincich [19], Chatterjee and Hadi [20], and Montgomery et al. [21], among others.

3.1. Regression Analysis: deaths versus BMI, death rate, life expectancy, population

In this sub-section, we provide the Minitab regression computer program outputs, followed by the interpretation of the results.

3.1.1. Minitab Regression Computer Program Output

The following are the Minitab regression computer program outputs based on our proposed multiple linear regression model, Equation (2.2).

(a) Regression Equation: It is given by

deaths = 472 - 14.4 BMI + 81.3 death rate - 1.00 life expectancy

(2.3) **- 0.00000033 population.**

(b) Analysis of Variance: It is provided in **Table 3.1** below.

Table 3.1

Analysis of Variance					
Source	DF	SS	MS	F	P
Regression	4	20173.0	5043.2	5.67	0.002
Residual Error	31	27561.0	889.1		
Total	35	47734.0			

(c) Model Summary: It is provided in **Table 3.2** below.

Table 3.2

s	R-Sq	R-Sq(adj)	Durbin-Watson Statistic
29.8172	42.3%	34.8%	2.53177

(d) Regression Coefficients: These are provided in **Table 3.3** below.

Table 3.3

Predictor	Coef	SE Coef	*t*	P	VIF
Constant	472.08	95.22	4.96	0.000	
BMI	-14.390	4.459	-3.23	0.003	1.9
death rate	81.30	27.16	2.99	0.005	1.3
life expectancy	-1.0011	0.7471	-1.34	0.190	1.6
population	-0.00000033	0.00000028	-1.15	0.257	1.5

3.1.2. Interpretation of the Results

In the following paragraphs, the key values in the Minitab regression computer program outputs (**Tables3.1 - 3.3**) are interpreted.

- From the Analysis of Variance, **Table 3.1**, we observe that the p-value is (**0.002**). This implies that that the model estimated by the regression procedure is significant at an α -level of 0.05. Thus at least one of the regression coefficients is different from zero.
- **The R^2 and Adjusted R^2 Statistic:** There are several useful criteria for measuring the goodness of fit of the multiple regression model. One such criteria is to determine the square of the multiple correlation coefficient R^2 (also called the coefficient of multiple determination), (see, for example, Draper and Smith [16], and Mendenhall and Sincich [18], among others). The R^2 value in the regression output (**Table 3.2**) indicates that 42.3% of the total variation of the response variable death (Y) values about their mean can be explained by the predictor variables used in the model. The adjusted R^2 value (or $R_a{}^2$) indicates that 34.8% of the total variation of the response variable death (Y) values about their mean can be explained by the predictor variables used in the model. As the values of R^2 and $R_a{}^2$ are not very different, it appears that at least one of the predictor variables contributes information for the prediction of Y. Thus, both values indicate that the model fits the data well.
- **Durbin-Watson Statistic:** From the **Table 3.2**, we observe that the Durbin-Watson statistic = 2.53177 > 2, which implies that the residuals are negatively correlated.
- From the **Table 3.3**, we observe that the p-values for the estimated coefficients of the mean BMI (X_1) and the death rate (X_2) are respectively 0.003 and 0.005, indicating that they are significantly related to the response variable death (Y) at an α -level of 0.05 significance. From the Table 3.3, we also observe that the p-values for the life expectancy (X_3) and the population (X_4) are relatively high, indicating that these are probably not related to the response variable death (Y) at an α -level of 0.05 significance.
- **Interpretation of $\mathbf{s}$, the Estimated Standard Deviation of $\boldsymbol{\varepsilon}$**: For our problem, from Table 3.2, we have $s = 29.8172$. Examination of this statistics indicates that approximately 95% of the actual mean death levels will fall within $2s = 59.6344$ of the values predicted by the first-order model, Equation (2.3).
- **Unusual Observations:** We note from the following **Table 3.4** (**Unusual Observations**) that the observations 1 and 36 are identified as unusual because the absolute values of the residuals are extremely large. This indicates that they are outliers.

Table 3.4

Unusual Observations						
Obs	**BMI**	**deaths**	**Fit**	**SE Fit**	**Residual**	**St Resid**
1	**27.5**	**84.00**	**95.20**	**29.19**	**-11.20**	**-1.85 X X**
36	**21.6**	**35.00**	**53.68**	**20.88**	**-18.68**	**-0.88 X**
X denotes a point that is an outlier in the predictors. **XX denotes a point that is an extreme outlier in the predictors.**						

- **Multicollinearity:** By multicollinearity, we mean that some predictor variables are correlated with other predictors. Various techniques have been developed to identify predictor variables that are highly collinear, and for possible solutions to the problem of multicollinearity, see, for example, Neter et al. [15], Draper and Smith [16], Tamhane and Dunlop [17], Mendenhall and Sincich [19], Chatterjee and Hadi [20], Montgomery et al. [21], and Vittinghoff et al. [30], among others. For example, we can examine the variance inflation factors (VIF), which measure how much the variance of an estimated regression coefficient increases if the predictor variables are correlated. Following Montgomery et al. [21], if the VIF is 5 - 10, the regression coefficients are poorly estimated. However, it has been observed by many researchers that for a large sample size, multicollinearity is not a big problem when compared to a small sample size. From **Table 3.3**, we observe that, since the variance inflation factors (VIF) for each of the estimated regression coefficient in our calculations are less than 5 for the mean BMI (X_1), death rate (X_2), life expectancy (X_3), and population (X_4), there does not seem to be multicollinearity for these predictors in our model.

- **Best Subsets Regression:** Another important criterion function for assessing the predictive ability of a multiple linear regression model is to examine the associated Mallows' C_p -statistic, including R-Sq (R^2), the percentage of variation in the response that is explained by the model, Adjusted R^2 (that is, R-Sq(adj), the percentage of the variation in the response that is explained by t for the number of predictors in the model relative to the number of observations), and S, the standard error of the estimate. The best subsets regression method is used to choose a subset of predictor variables so that the corresponding fitted regression model optimizes the Mallows' C_p -statistic, which may be interpreted as follows:
  - ✓ A Mallows' Cp value that is close to the number of predictors plus the constant model produces relatively precise and unbiased estimates.
  - ✓ A Mallows' Cp value that is greater than the number of predictors plus the constant model is biased and does not fit the data well.

Based on the above-mentioned criteria, the following (**Table 3.5**) gives the possible predictor model.

Table 3.5

	Vars	R-Sq	R-Sq(adj)	C-p	S	Possible Predictor Model
(i)	4	42.3	34.8	5.0	29.817	In the presence of death rate (X_2), life expectancy (X_3), and population (X_4), BMI (X_1) is a good predictor of response variable death (Y).
(ii)	4	42.3	34.8	5.0	29.817	In the presence of BMI (X_1), life expectancy (X_3), and population (X_4), death rate (X_2) is a good predictor of response variable death (Y).

4. ANALYSIS OF THE MODEL BASED ON THE RESIDUAL PLOTS OF THE DEATH (Y)

The Minitab Version 17.0 regression computer program outputs for residual plots of Y are given in Figures 4.1 – 4.3 below. In the paragraphs that follow, the residuals of the regression model (2.3) are analyzed using the graphs in Figures 4.1 – 4.3, and examine the goodness of fit of the model based on these residual plots.

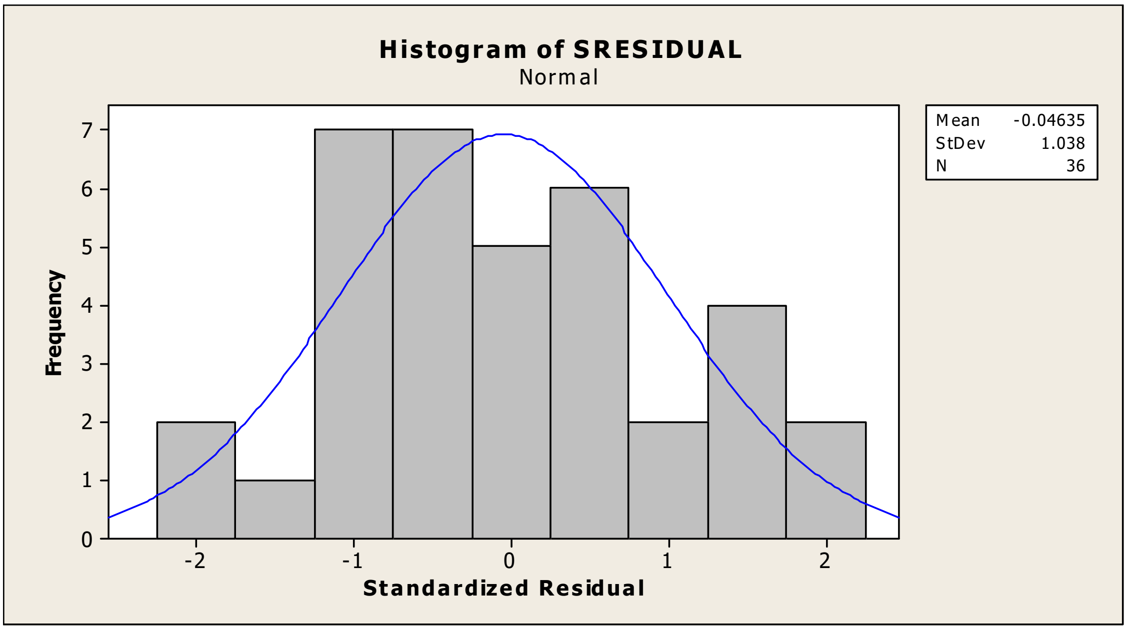

Figure 4.1. MINITAB Residual Graph for Model (2.3)

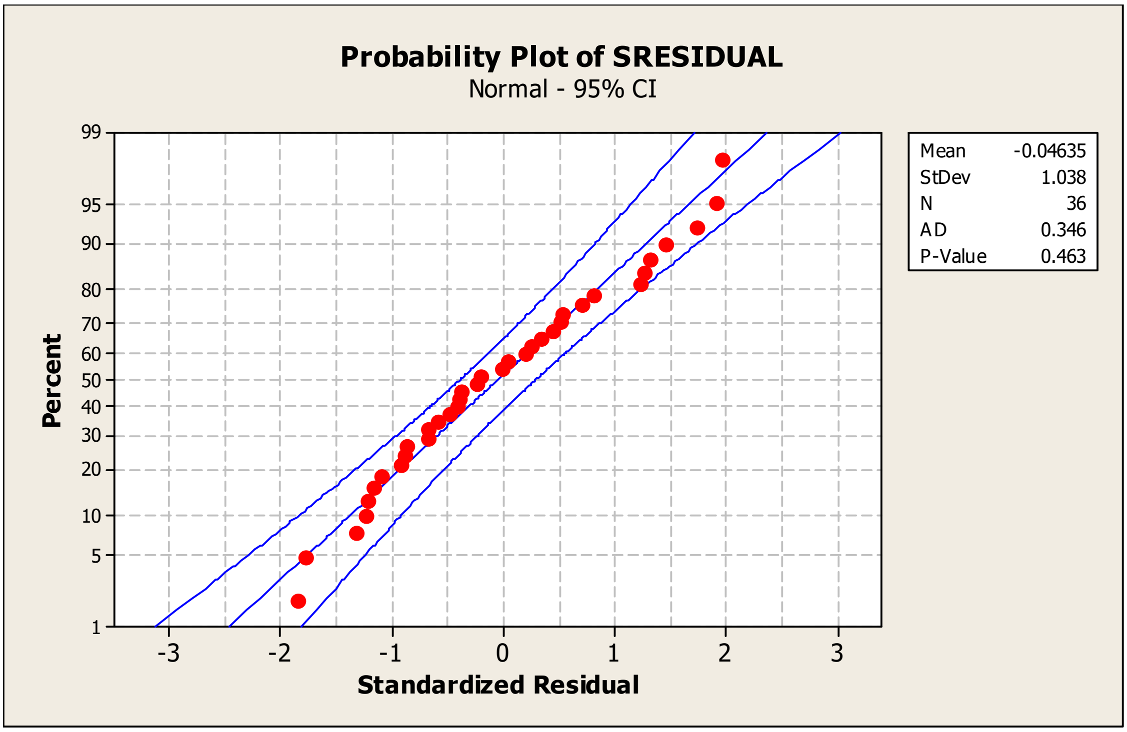

Figure 4.2. MINITAB Residual Graph for Model (2.3)

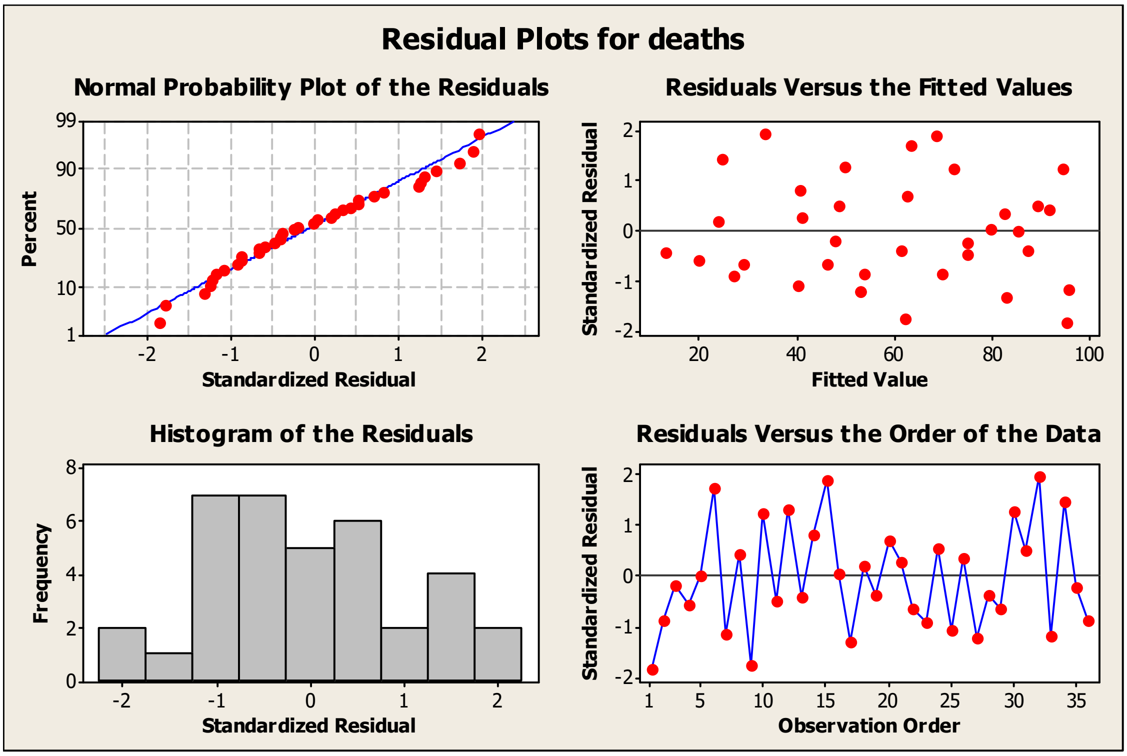

Figure 4.3. MINITAB Residual Graphs for Model (2.3)

4.1. Interpretation of the Graphs (Figures 4.1 – 4.3)

- The Minitab printout in Figures 4.1 – 4.2 shows both a histogram and a normal probability plot for the standardized residuals. It is obvious from these graphs that the regression errors appear to be normal, except one or two outliers.
- From the normal probability plots in Figures 4.2 – 4.3, we observe that there exists an approximately linear pattern. This indicates the consistency of the data with a normal distribution. The outliers are indicated by the points in the upper-right and left-bottom corners of the plot.
- From the plot of residuals versus the fitted values in Figure 4.3, it is evident that the residuals get smaller, that is, closer to the reference line, as the fitted values increase. This may indicate that the residuals have non-constant variance (see, for example, Draper and Smith [16], among others, for details).
- The histogram of the residuals in Figure 4.3 indicates that no outliers exist in the data.
- The plot for residuals versus order is also provided in Figure 4.3. It is defined as a plot of all residuals in the order that the data was collected. It is used to find non-random error, especially of time-related effects. A clustering of residuals with the same sign indicates a positive correlation, whereas a negative correlation is indicated by rapid changes in the signs of consecutive residuals.

5. CONCLUDING REMARKS

From the above analysis, it appears that our multiple regression model developed to analyze the impact of the body mass index and other factors such as the death rate, life expectancy, and population on the mortality (death), Y, and for predicting it, based on a sample data of some countries affected during Covid-19, is useful and adequate. In the presence of $\boldsymbol{X_2}, \boldsymbol{X_3}$, and $\boldsymbol{X_4}$, X_1 is a good predictor of Y. In the presence of $\boldsymbol{X_1}, \boldsymbol{X_3}$, and $\boldsymbol{X_4}$, X_2 is a good predictor of Y. As the values of R^2 and $R_a{}^2$ are not very different, it appears that at least one of the predictor variables contributes information for the prediction of Y. Thus, our multiple regression model for predicting the mortality (death), Y, seems to be useful and adequate, and the overall regression is statistically significant. For future work, one can consider to develop and study similar models for other issues and problems associated with the fields of medical, biological, behavioral, and other applied sciences. One can also develop similar models by adding other variables, for example, the gender, marital status, employment status, race and ethnicity of the adults, as well as the squares, cubes, and cross products of $\boldsymbol{X_1}, \boldsymbol{X_2}, X_3$, **and** $\boldsymbol{X_4}$**.** In addition, one could also study the effect of some data transformations. We believe that the present study would be useful for researchers in the fields of medical and other applied sciences.

Authors' Contributions: All authors have equally made contributions. Also, all authors read and approved the final manuscript.
Funding: The authors state that they have no funding source for this paper. Further, the authors state that no funding source or sponsor has participated in the realization of this work.
Availability of Data and Materials: Not applicable.
Declarations Conflict of interest: The authors declare that they have no competing interests.

APPENDIX 1

(Source: Gabbrielli R., and Pugno N. M. (2023) [14])

Country	Start Date	Deaths	Life Expectancy	Population	Death Rate	BMI
Andorra	3/2/2020	84	83.73	79843	1.26317	27.5
Benin	3/16/2020	44	61.77	13352864	0.00415	23.4
Botswana	3/30/2020	42	69.59	2630300	0.02112	24.7
Brunei	3/9/2020	3	75.86	449002	0.00821	26.2
Burkina Faso	3/10/2020	85	61.58	22673764	0.00462	22.1
Cape Verde	3/20/2020	113	72.98	593162	0.24313	24.7
Central African Republic	3/15/2020	63	53.28	5579148	0.01416	22.4
Chad	3/19/2020	104	54.24	17723312	0.00746	22.3
Comoros	4/30/2020	10	64.32	836783	0.0178	24.1
Congo	3/15/2020	108	64.57	5970430	0.02269	23.3
Djibouti	3/18/2020	61	67.11	1120851	0.06897	23.3
Equatorial Guinea	3/15/2020	86	58.74	1674916	0.0644	25.6
Fiji	1/30/2020	2	67.44	929769	0.00234	27.2
Guinea	3/13/2020	64	66.47	2388997	0.03349	25.5
Guinea-Bissau	3/25/2020	124	62.05	2705995	0.05788	24
Guinea	3/13/2020	81	61.6	13859349	0.00728	22.7
Guinea-Bissau	3/25/2020	45	58.32	2105580	0.02776	23.1
Iceland	2/28/2020	29	82.99	372903	0.09246	25.9
Lesotho	5/13/2020	51	54.33	2305826	0.0348	24.9
Liberia	3/17/2020	83	64.1	5302690	0.01977	24
Maldives	3/8/2020	48	78.92	523798	0.11224	25.1
Mauritius	3/18/2020	10	74.99	1299478	0.00975	25.6
Mongolia	3/10/2020	1	69.87	3398373	0.00036	26
Niger	3/20/2020	104	62.42	26207982	0.00506	21.7
Papua New Guinea	3/20/2020	9	64.5	10142625	0.00113	25.3
Rwanda	3/14/2020	92	69.02	13776702	0.00835	22

Sao Tome and Principe	**4/6/2020**	**17**	**70.39**	**227393**	**0.10144**	**24.8**
Sierra Leone	**3/31/2020**	**76**	**54.7**	**8605723**	**0.01172**	**22.8**
Singapore	**1/23/2020**	**29**	**83.62**	**5637022**	**0.00548**	**23.7**
Somalia	**3/16/2020**	**130**	**57.4**	**17597508**	**0.0093**	**21.9**
South Sudan	**4/5/2020**	**63**	**57.85**	**10913172**	**0.0078**	**25.2**
Tajikistan	**5/1/2020**	**90**	**71.1**	**9952789**	**0.01353**	**25.4**
Tanzania	**3/16/2020**	**21**	**65.46**	**65497752**	**0.0004**	**23.1**
Thailand	**1/4/2020**	**61**	**77.15**	**71697024**	**0.00086**	**24.1**
Togo	**3/4/2020**	**68**	**61.04**	**8848700**	**0.00929**	**23.2**
Vietnam	**1/23/2020**	**35**	**75.4**	**98186856**	**0.00038**	**21.6**

References

[1] Zhou F, Yu Tu, Du R, Fan G, et.al., Clinical Course and Risk Factors for Mortality of Adult Inpatients With COVID-19 in Wuhan, China: A Retrospective Cohort Study, Lancet (2020) 395(10229): 1054-62. Doi:10.1016/S0140-6736(20)30566-3.

[2] Organization WH. Naming the Coronavirus Disease (COVID-19) and the Virus That Causes it (2020). World Health Organization. Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(COVID-19)- and-the-virus-that-causes-it (Accessed March 17, 2021).

[3] World Health Organization. Director- General's Remarks at the Media Briefing on 2019-nCoV on 11 February 2020. http://www.who.int/dg/speeches/detail/who-director-generals-remark-at-the-media-briefing-on-2019-ncov-on 11-february-2020 (Accessed on February 12, 2020).

[4] Obesity: Preventing and Managing the Global Epidemic, Report of a WHO Consultation, World Health Organ Tech Rep Ser.2000, 894: i-xii, 1-253.

[5] Jia, P., Obesogenic environment and Childhood Obesity. Obes Rev (2021). http://doi.org/10.1111/obr.13158

[6] Martin, V., Castilla, J. Godoy P., et.al., High body mass index as a Risk Factor for Hospitalization Due to Influenza: A Case-Control Study, Arch Bronconemol, (2016), 52, 299-307.

[7] Badawi, A., Ryoo SG., Prevalence of Diabetes in the 2009 Influenza A (H1N1) and the Middle East respiratory syndrome coronavirus: a systematic review and Meta-analysis, J Public Health Res, 2016, 5,733.

[8] Petrilli, CM., Jones, SA., Yang J, et.al., Factors Associated with Hospital Admission and critical illness among 5279 People with Coronavirus Disease 2019 in New York City: prospective cohort study, BMJ,2020, 369:m1966. 10.1136/bmj.m1966.

[9] Patel, A., Jernigan, DB., nCo VCDCRT. Initial public health response and interim clinical guidance for the 2019 novel coronavirus outbreak - United States, December 31, 2019-February 4, 2020. MMWR Morb Mortal Wkly Rep 2020; 69:140-6.

[10] Kompaniyes, L, Goodman, AB., Belay B. et.al., Body Mass Index and Risk for COVID-19-Related Hospitalization, Intensive Care Unit Admission, Invasive Mechanical Ventilation, and Death- United States, March-December 2020, (2021) MMWR Morb Mortal Wkly Rep 2021, 70, 355-361. Doi: http://dx.doi.org/10.15585/mmwr.mm7010e4

[11] Habib, Y., Aislim, R., Aurbidi, L., et.al., Effect of obesity on clinical outcomes in COVID-19 patients, Cureus -Jan 13-2023.

[12] Palaiodimos, I., Kokkinidis, DG., Li, W., et.al. Severe obesity, increasing age and male sex are independently associated worse in-hospital outcomes and higher in-hospital mortality in a cohort of patients with COVID-19 in the Bronx, New York, Metabolism 2020; 108:154262.

[13] Popkin, BM., Du, S., Green, WD., Beck, MA., et.al., Obesity as a risk factor for COVID-19 mortality in women and men in the UK biobank: Comparisons with influenza/pneumonia and coronary heart disease, Diabetes, Obes Metab. (2021), 23, 258-62. Doi: 10.1111/dom.14199.

[14] Gabbrielli, R., Pugno, NM., The impact of mean body mass index on reported mortality from COVID-19 across 181 countries. Front. Public Health, 2023, 11: 1106313. Doi: 10.3389ppubh. 2023.1106313.

[15] Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W., *Applied Linear Statistical Models, 4th Edition,* WCB McGraw-Hill, New York, USA, (1996).

[16] Draper, N. R., and Smith, H., *Applied Regression Analysis (3rd edition)*, John Wiley & Sons, INC., New York, USA, (1998).

[17] Tamhane, A. C., and Dunlop, D. D., *Statistics and Data Analysis: From Elementary to Intermediate (1st edition)*, Pearson Prentice Hall, Upper Saddle River, NJ, USA, (2000).

[18] Mendenhall, W., and Sincich, T., *A Second Course in Statistics: Regression Analysis (8th edition),* Pearson-Prentice Hall, Upper Saddle River, NJ, USA, (2023).

[19] Chatterjee, S., and Hadi, A. S., *Regression Analysis by Example, 5th Edition*. John Wiley & Sons, New York, USA, (2012).

[20] Montgomery, D. C., Peck, E. A., and Vining, G. G., *Introduction to Linear Regression Analysis, Fourth Edition,* John Wiley & Sons, New York, USA, (2013).

[21] Suárez, E., Perez, C. M., Rivera, R., and Martinez, M. N., *Applications of Regression Models in Epidemiology*, John Wiley & Sons, New York, USA, (2017).

[22] Cleophas, T. J., and Zwinderman, A. H., *Regression Analysis in Medical Research for Starters and 2nd Levelers*, Springer, Basel, Switzerland, (2018).

[23] Guzman, C. I., and Kibria, B. M. G., Developing Multiple Linear Regression Models for the Number of Citations: A Case Study of Florida International University Professors, *International Journal of Statistics and Reliability Engineering*, (2019), **6(2)**, 75-81.

[24] Johnson, R. A., and Wichern, D. W., *Applied Multivariate Statistical Analysis, 9th Edition*. Pearson-Prentice Hall, Upper Saddle River, NJ, USA, (2023).

[25] Kamel, A., and Abonazel, M. R., A Simple Introduction to Regression Modeling using R, *Computational Journal of Mathematical and Statistical Sciences*, (2023), *2*(1), 52-79. doi: 10.21608/cjmss.2023.189834.1002.

[26] Kibria, B. M. G., Linear and Non-linear Regression Models: Theory, Simulation and Applications, *24th ANNUAL CONFERENCE of VIJNANA PARISHAD OF INDIA and NATIONAL SEMINAR on NEW THRUST AREAS IN MATHEMATICS, MATHEMATICAL SCIENCES, AND ENGINEERING*, INDIA, April 28-30, (2023).

[27] Siegrist, M., and Kibria, B. M. G., Predicting Total Number of Deaths using COVID-19 World Data: Application of Linear Regression Model. *Int J Ind Operations Res*, (2020), 3:008.

[28] Ghosal, S., Sengupta, S., Majumder, M., and Sinha, B., Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020), *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, (2020), 14 (4), 311-315. https://doi.org/10.1016/j.dsx.2020.03.017

[29] Kibria, B. M. G., and Urbistondo, M., Predicting the Total number of COVID-19 Deaths by County, *WSEAS TRANSACTIONS on ENVIRONMENT and DEVELOPMENT*, (2021), 17, 1299 – 1310. DOI: 10.37394/232015.2021.17.119.

[30] Vittinghoff, E., Glidden, D. V., Shiboski, S. C., and McCulloch, C. E., *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models, Second Edition*, Springer New York, NY, USA (2014).

[31] Trefethen, N.,"New BMI (Body Mass Index)", A letter to the editor, 5th January 2013 Issue, the *Economist*, London, UK (2013). https://people.maths.ox.ac.uk/trefethen/bmi.html.

[32] https://medlineplus.gov/ency/article/007196.htm: “Body mass index”. MedlinePlus, National Library of Medicine (NLM), National Institutes of Health (NIH), MD, USA.