Journal of Mathematics and Statistics Studies
ISSN: 2709-4200
DOI: 10.32996/jmss
Journal Homepage: www.al-kindipublisher.com/index.php/jmss

JMSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

A Statistical Analysis of Positive Excess Mortality at Covid-19 in 2020-2021

Md Nurul Raihen[1] ✉ Sultana Akter[2], Fariha Tabassum[3], Farjana Jahan[4] and Md Nazmul Sardar[5]

[1]*Assistant Professor, Department of Mathematics and Computer Science, Fontbonne University, St. Louis, MO, 63105, USA*
[2]*Teaching Assistant, MS in Statistics, Department of Statistics, Western Michigan University, Kalamazoo, MI, 49006, USA*
[3]*Teaching Assistant, MS in Sociology, Department of Sociology, Western Michigan University, Kalamazoo, MI, 49006, USA*
[4]*Teaching Assistant, MS in Statistics, Department of Statistics, Western Michigan University, Kalamazoo, MI, 49006, USA*
[5]*Senior Officer, Product Development at Radiant Nutraceuticals Ltd (F. Hoffman-La Roche Ltd), Dhaka, 1000, Bangladesh*

Corresponding Author: Md Nurul Raihen, **E-mail**: fy4233@wayne.edu

| ABSTRACT

When it comes to making assessments about public health, the mortality rate is a very important factor. The COVID-19 pandemic has exacerbated well-known biases that affect the measurement of mortality, which varies with time and place. The COVID-19 pandemic took the world off surveillance, and since the outbreak, it has caused damage that many would have thought unthinkable in the present era. By estimating excess mortality for 2020 and 2021, we provide a thorough and consistent evaluation of the COVID-19 pandemic's effects. Excess mortality is a term used in epidemiology and public health to describe the number of fatalities from all causes during a crisis that exceeds what would be expected under 'normal' circumstances. Excess mortality has been used for thousands of years to estimate health emergencies and pandemics like the 1918 "Spanish Flu"6. Positive excess mortality occurs when actual deaths exceed previous data or recognized patterns. It could demonstrate how a pandemic affects the mortality rate. The estimates of positive excess mortality presented in this research are generated using the procedure, data, and methods described in detail in the Methods section and briefly summarized in this study. We explored different regression models in order to find the most effective factor for our estimates. We predict the pandemic period all-cause deaths in locations lacking complete reported data using the Poisson, Negative Binomial count framework. By overdispersion test, we checked the assumption of the Poisson model, and then we chose the negative binomial as a good fitting model for this analysis through Akaike Information Criteria (AIC) and Standardized residual plots, after that checking the P-value<0.05; we found some significant predictors from our choosing model Negative binomial model, and the coefficient of all predictors gave the information that some factors have a positive effect, and some has a negative effect at positive excess mortality at COVID-19 (2020-2021).

| KEYWORDS

COVID-19, Excess Mortality, Pandemic, Poisson Regression, Negative Binomial regression, WHO (World Health Organization), Region

| ARTICLE INFORMATION

ACCEPTED: 24 July 2023 **PUBLISHED:** 03 August 2023 **DOI:** 10.32996/jmss.2023.4.3.2

1. Introduction

Coronaviruses are a type of virus. A coronavirus identified in 2019, SARS-CoV-2, has caused a pandemic of respiratory illness called COVID-19. The World Health Organization has been tracking the impact of COVID-19 as the pandemic has evolved over time. The pandemic's conditions, which have overburdened some health systems and caused some patients to avoid care, have caused further deaths. Due to COVID-19 data difficulties, excess mortality is a more objective and comparable indicator. The WHO defines excess mortality as "the mortality above what would be expected based on the non-crisis mortality rate in the population of interest" (World Health Organization, 2023).

The excess mortality associated with the COVID-19 pandemic is utilized to quantify the pandemic's direct and indirect effects. Excess mortality is the difference between the predicted total number of deaths for a particular location and time period and the number that would have been expected in the absence of a crisis (e.g., the COVID-19 pandemic). In accordance with the difficulties associated with using reported data on COVID-19 cases and fatalities, excess mortality is viewed as a more objective and comparable measure that accounts for both the direct and indirect effects of the pandemic (World Health Organization, 2023).

The World Health Organization is being updated with overall case and mortality rates. These figures obviously don't give a clear picture of the COVID-19-related health burden or the number of lives lost as a result of the epidemic. Some of the fatalities that can be attributed to COVID-19 have not been confirmed as being caused by the virus because premortem tests were not carried out. There has been a wide range of approaches to death certification across countries, especially in the face of co-morbidities and the emergence of COVID-19 (Raihen et al., 2023).

The primary outcome to evaluate excess death was excess mortality during the pandemic, defined as the difference between the number of reported all-cause deaths and the expected number of deaths during the pandemic. The calculation formula was excess mortality = estimated deaths−expected deaths population. What is the meaning of positive excess mortality? Positive excess mortality occurs when the number of expected deaths from prehistoric pandemic data is less than the number of deaths observed during the Covid-19 period. This positive excess mortality can be caused by several factors, including the seriousness of the disease, the burden on the healthcare system, changes in behavior or social connections, and other indirect effects.

1.1 Research Questions

- The major goal of this study is to determine which factors (region, year, sex, age_ group, and pop) are significant with different aspects of excess mortality presence of Covid-19.
- In order to put this idea, various quantitative research strategies were used and figured out which regression model was best fit for the positive excess mortality mean considering the people of COVID-19 disease.

2. Literature Review

According to the World Health Organization (WHO), there were an estimated 14.9 million excess deaths caused by COVID-19 worldwide between 2020 and 2021 (Jha, P. et al., 2022). The data provided by WHO for deaths worldwide are lower than the estimates provided by the Institute for Health Metrics and Evaluation (IHME)2, which reported 18.2 million deaths (17.1 million–19.6 million), and the estimates provided by The Economist, which provided 17.7 million deaths (13.9 million–21.1 million) for the same time period. On the other hand, estimates from the government based on data from the Coronavirus App on the number of deaths caused by COVID-19 over the world in 2020–2021 imply that the number is less than 6 million (Rocco, P. et al., 2021).

When making public health decisions, mortality rates are crucial (Collaborators, C. E. M, 2022). On the other hand, countries, health systems, and individual physicians all classify deaths differently (Karlinsky, A. 2021, Ramírez-Soto, M. C. et al., 2022). Inadequate tests and overloaded health systems, caused by a sudden increase in COVID-19 symptom patients in most countries, may have led to an underestimation of COVID-19-related fatalities in the early stages of the pandemic (Achilleos, S. et al., 2022).

Furthermore, deaths that are indirect during the pandemic, such as those caused by resource restrictions in health care systems, unnatural causes, or severe occurrences, are likely to be misclassified as direct mortality of COVID-19 due to misdiagnosis and pandemic "bias" (C. E. M, 2022, Ioannidis, Ramírez-Soto, M. C., et al., 2022, J. P. et al., 2021). Even Prior to the massive reorganization of death causes caused by COVID-19, death certificates already had a bad reputation for being inaccurate. COVID-19 and other diagnoses on the death certificate may be more difficult to assign if comorbidities are present (Gobiņa I. et al., 2022, Kiang M. V. et al., 2020, Koffman J. et al., 2020).

Death rates, death counts, and life expectancy are only a few of the metrics that can be used to provide an overview of mortality in a given location. Death rates and death counts are the most widely used indices for determining excess mortality (Bilinski & Emanuel, 2020; Faust et al., 2020; Schöley, 2021). Researchers have used a wide variety of death rates, including crude death rates (CDRs) (Aburto et al., 2021; Basellini et al., 2021; Németh et al., 2021; Stokes et al., 2021) and age-specific death rates (Németh et al., 2021) and age-standardized death dates (SDRs) (Islam et al., 2021; Krieger et al., 2020) Variation in the predicted mortality level used to assess excess mortality may emerge from the fact that these indices capture a wide range of mortality levels and trends.

In this paper, we introduce the World Death Dataset, an effort to collect and maintain global death statistics that are updated on a regular basis (Raihen, M. N., & Akter, S. et al., 2023). The dataset is publicly available and is almost daily updated at https://www.who.int/data/sets/global-excess-deaths-associated-with-covid-19-modelled-estimates. Since the work of our research, the dataset has been added to Our World in Data (Giattino et al., 2020) and The Economist's and Financial Times' excess

mortality trackers. We conclude that data from all different locations is reliable enough to permit computation of excess mortality (see Discussion), while not all countries give data of the same quality or detail.

In addition to deaths caused by COVID-19 infection, it is possible that social distancing mandates and other pandemic restrictions reduced the number of deaths caused by certain diseases and injuries, such as those caused by traffic accidents (Mungmunpuntipantip, R. et al., 2021, Rozenfeld, M. et al., 2021, Salottolo, K. et al., 2021), while increasing the number of deaths caused by others, such as those caused by chronic and acute conditions affected by delayed care-seeking in overstretched health-care systems (Zubiri, L. et al., 2021, Folino, A. F. et al., 2020) in comparison to expected or baseline conditions. It is difficult to determine how much of the excess mortality is attributable to COVID-19 infection and how much is attributable to other societal, economic, or behavioral changes associated with the pandemic. This is especially the case because there is a lack of detailed data on the specific causes of death in many countries. Understanding the total mortality impact that the pandemic has is a key first step, even though it will be extremely necessary to identify the factors that contributed to the excess mortality that was observed.

3. Methodology

3.1 Data Source

As the COVID-19 epidemic has grown over time, the World Health Organization (WHO) has monitored its effects. WHO (World Health Organization) has been informed how many people have been infected and how many have died. Besides the deaths that can be directly linked to the pandemic, there are also deaths that can be linked to conditions that have been around since the pandemic started and have caused some health systems to be overloaded or some patients to avoid getting care. Since using reported COVID-19 data can be hard, excess mortality is thought to be a more objective and similar measure. The World Health Organization (WHO) says that "excess mortality" is a death rate that is higher than what would be predicted based on the normal death rate in the population of interest. According to data collection history, the official WHO estimate published the excess mortality data expected as yearly based frequencies on May 20, 2021.

3.2 Data Structures

Death Data is always counted, and the data on excess mortality is also counted, and its spatial coverage is Global. In this article, we will detail the various models we performed on this dataset, as well as the methods we used to draw conclusions about the excess mortality of COVID-19 and to predict the efficient predictors. In order to properly conduct any data analysis or data operation, we must first possess extensive background knowledge in the field in question. Therefore, we will discuss data set characteristics and how they relate to one another. Five of the attributes in this data collection are classified, and three are numerical attributes. In the latest estimated data at WHO, there are 224 observations collected about the excess deaths associated with the COVID-19 pandemic from all causes by age, sex, and year.

Table 1 List of features and their description in the initial dataset (the dataset is also available at the WHO website (Global excess deaths associated with COVID-19 (modelled estimates) (who.int))

Feature name	Type	Description and values	% missing
location	Nominal	Region (Global or WHO region), AFR: African Region; AMR: Region of the Americas; EMR: Eastern Mediterranean Region; EUR: European Region; WPR: Western Pacific Region; SEAR: South-East Asia Region	0%
year	Nominal	year of death (2020 and 2021)	0%
sex	Nominal	gender (male and female)	0%
age	Nominal	age- group from 0 to 85+ (0-24,25-34,35-44,45-54,55-64,65-74,75-84,85+)	0%
Pop	Numeric	Sex-and age-specific population number	0%
type	Nominal	estimated type for select year (reported or predicted)	0%
expected. mean	Numeric	expected deaths from all-causes by age, sex and year (mean)	0%

acm. mean	Numeric	estimated deaths from all-causes by age, sex, and year (mean)	0%
Excess. mean.	Numeric	excess deaths associated with COVID-19 pandemic from all-causes by age, sex and year (mean)	0%

3.2 Data Manipulation

There is a total of 224 data points in the WHO collection (word.csv data); perhaps here we got 184 observations, which contain only positive excess mortality. For our analysis perspective, we considered our explanatory variables to be location, year, age group, sex, and pop, and the response variable to be excess.mean (positive excess mortality).

3.3 Model Procedure

We use the programming language R to fit Poisson regression and Negative Binomial models to the Mortality data because the data is of the count type. Excel, Microsoft Word, and R-code have all been put to use at various points throughout our study to perform manipulations on raw data. We provide an outline of the theory as well as its implementation in R (R Development Core Team 2008) for certain fundamental count data regression models such as Poisson and negative binomial (see **Table 2** for an overview of the model's components).

They are implemented in R by the glm() function (Chambers and Hastie 1992) in the stats package and the glm.nb() function in the MASS package (Venables and Ripley 2002), and negative binomial regression is extended of GLM. The classical Poisson model and the negative binomial model are both described in a generalized linear model (GLM) framework. Objects should be modeled so that they are more comparable to their glm() and glm.nb() counterparts. Although all GLMs utilize the same log-linear mean function ($\log(\mu) = x > \beta$), each GLM makes a unique set of assumptions regarding the likelihood of the remaining variables (Raihen, M. N., Akter, S., & Sardar, M. N., et al., 2023).

3.4 Analysis

In this section, we discuss the data processing, data summary, and methodology used to get estimates of increased mortality based only on the presence of Covid-19. In this investigation, we use both direct and indirect COVID-19 death. Estimating the number of direct and indirect deaths caused by COVID-19 required the Department of Economic and Social Affairs of the United Nations (UN DESA) and the WHO Regulations for Scientific and Advisory Groups to work together to produce harmonized excess mortality procedures (Stein Jr, I., & Raihen, M. N., et al., 2023).

Table 2 Summary of the Data for explanatory and response predictors

```
   location             year          sex
 Length:184         Min.   :2020   Length:184
 Class :character   1st Qu.:2020   Class :character
 Mode  :character   Median :2021   Mode  :character
                    Mean   :2021
                    3rd Qu.:2021
                    Max.   :2021
  age_group             pop               type
 Length:184         Min.   :5.174e+05   Length:184
 Class :character   1st Qu.:1.998e+07   Class :character
 Mode  :character   Median :5.552e+07   Mode  :character
                    Mean   :1.166e+08
                    3rd Qu.:1.152e+08
                    Max.   :1.663e+09
  excess.mean.
 Min.   :    362
 1st Qu.:  25877
 Median :  64830
 Mean   : 164936
 3rd Qu.: 179780
 Max.   :1450563
```

3.5 The Poisson Regression Model

Since mortality data is count type, we fitted our data with the generalized linear model (glm) of the Poisson regression family.

Model 1

Table 3: Coefficient Table for the Poisson regression model

```
Coefficients:
                Estimate Std. Error  z value Pr(>|z|)
(Intercept)    -1.560e+03  7.952e-01 -1961.33   <2e-16 ***
locationAMR     1.256e+00  1.188e-03  1057.43   <2e-16 ***
locationEMR     3.567e-01  1.364e-03   261.45   <2e-16 ***
locationEUR     1.231e+00  1.192e-03  1032.91   <2e-16 ***
locationGlobal  2.753e+00  1.230e-03  2238.67   <2e-16 ***
locationSEAR    1.891e+00  1.129e-03  1675.42   <2e-16 ***
locationWPR    -2.384e-01  1.745e-03  -136.58   <2e-16 ***
year            7.771e-01  3.935e-04  1974.71   <2e-16 ***
sexMale         2.003e-01  3.665e-04   546.40   <2e-16 ***
age_group0-24  -4.632e+00  5.746e-03  -806.15   <2e-16 ***
age_group25-34 -2.361e+00  1.882e-03 -1254.63   <2e-16 ***
age_group35-44 -1.450e+00  1.244e-03 -1165.04   <2e-16 ***
age_group45-54 -5.374e-01  9.661e-04  -556.27   <2e-16 ***
age_group55-64  4.765e-02  7.800e-04    61.09   <2e-16 ***
age_group65-74  3.539e-01  6.370e-04   555.47   <2e-16 ***
age_group75-84  8.379e-02  6.041e-04   138.70   <2e-16 ***
pop             1.525e-10  2.674e-12    57.01   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The coefficient table of the Poission regression model shows that all predictors are significant at a p-value<0.05. The asterisks and dot denote the significance (different from zero) of each regression parameter as follows: No asterisk and no dot: NOT significant, no asterisk and a single dot: significant at the 10% level and one or more asterisks: significant at the 5% level.

Assumption Checking

1.

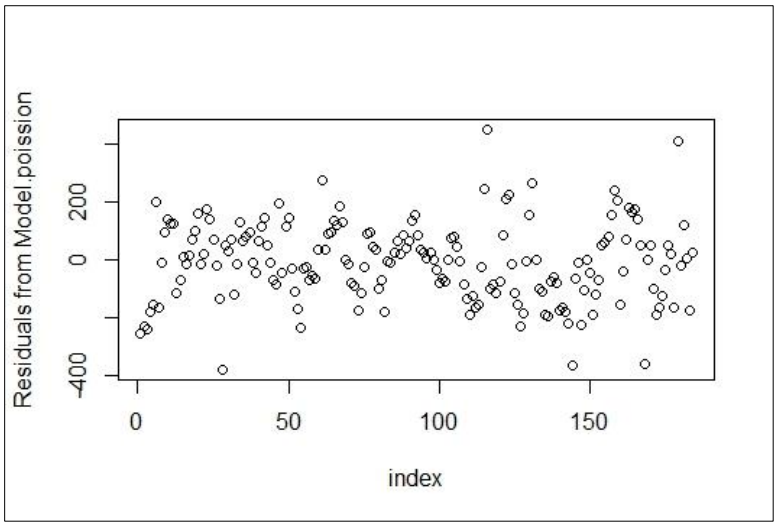

Figure 1: Residula plot of poisson regression

From the above figure, we can see from the residuals plot of fitting **model 1** (Poisson Regression) that the residuals have spread out in a random way with no properly clear pattern. This shows that the model fitting for model 1 has no systematic bias, which is a good sign for a model. But we must consider more things for assumption checking.

2.

The Poisson distribution implies that the mean and standard deviation are equal. Occasionally, our data exhibit excess variation that exceeds the mean. This condition is known as overdispersion. We also examine the overdispersion of our model to determine whether the variance value is greater than the mean value. If the variance value is greater than the mean value, the Poisson regression model's assumption that the mean and variance values are equal is violated. By dividing the residual deviation by the degrees of freedom, one can detect overdispersion.

Table 4: Overdispersion test for checking the Poisson model assumption.

```
        Overdispersion test

data:  model.poi
z = 7.0535, p-value = 8.726e-13
alternative hypothesis: true alpha is greater than 0
sample estimates:
  alpha
19447.5
```

p-value **< 0.05** indicates overdispersion exists.

Overdispersion can be detected by dividing the residual deviance by the degrees of freedom. If this quotient is much greater than one, the negative binomial distribution should be used, and the p-value if the p-value from the overdispersion test is less than 0.05, it means overdispersion exists.

3.6 The Negative Binomial Regression Model

A popular generalization of Poisson regression is negative binomial regression. The negative binomial distribution can be used as an alternative to the Poisson distribution for overly dispersed data. The Poisson distribution implies that the mean and variance are identical. Occasionally, data exhibit more variation that exceeds the mean. This condition is known as overdispersion, and negative binomial regression is more flexible than Poisson regression in this regard. Negative binomial models assume that the conditional means and variances are not equal. The negative binomial model has one parameter more than the Poisson regression that adjusts the variance independently from the mean. This parameter, often denoted as the dispersion parameter, allows the variance to exceed the mean, thus capturing overdispersion. Consequently, the Poisson model is nested within the negative binomial model (Charro, F., Ali, A. H., Raihen, N., Torres, M., & Wang, P. et al., 2023).

Model 2

Table 5: Coefficient table of negative binomial regression model

```
Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)    -1.587e+03  1.780e+02  -8.915  < 2e-16 ***
locationAMR     9.427e-01  1.597e-01   5.902 3.59e-09 ***
locationEMR     3.312e-02  1.577e-01   0.210 0.833637
locationEUR     6.571e-01  1.566e-01   4.198 2.70e-05 ***
locationGlobal  2.502e+00  2.047e-01  12.223  < 2e-16 ***
locationSEAR    1.357e+00  1.573e-01   8.627  < 2e-16 ***
locationWPR    -6.501e-01  1.881e-01  -3.457 0.000546 ***
year            7.908e-01  8.809e-02   8.977  < 2e-16 ***
age_group0-24  -3.497e+00  2.799e-01 -12.495  < 2e-16 ***
age_group25-34 -2.452e+00  2.020e-01 -12.141  < 2e-16 ***
age_group35-44 -1.556e+00  1.689e-01  -9.214  < 2e-16 ***
age_group45-54 -5.925e-01  1.634e-01  -3.627 0.000287 ***
age_group55-64 -9.698e-02  1.607e-01  -0.604 0.546120
age_group65-74  1.577e-01  1.584e-01   0.995 0.319546
age_group75-84 -9.865e-02  1.607e-01  -0.614 0.539202
sexMale         3.104e-01  8.594e-02   3.612 0.000304 ***
pop            -7.307e-10  3.719e-10  -1.965 0.049457 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From this result, we got some significant predictors for positive excess mortality during COVID-19.

The asterisks and dot denote the significance (different from zero) of each regression parameter as follows: No asterisk and no dot: NOT significant, No asterisk and a single dot: significant at 10% level and one or more asterisks: significant at 5% level (Mansfield, E. R. et al., 1982).

3.7 Comparison of the Poisson Regression Model and the Negative Binomial Regression Model

Table 6: Akaike Information Criteria (AIC) and Residual deviance

Model	df	AIC	Residual. Deviance	Comment
1.model.poisson	17	3351820.185	3,349,425	The AIC and residual. deviance value both are lower for model2 so negative binomial is more appropriate model
2.model.nb	18	4422.488	194	

A model with smaller AIC and smaller residual deviance is a better model (Yamaoka, K. et al., 1978).

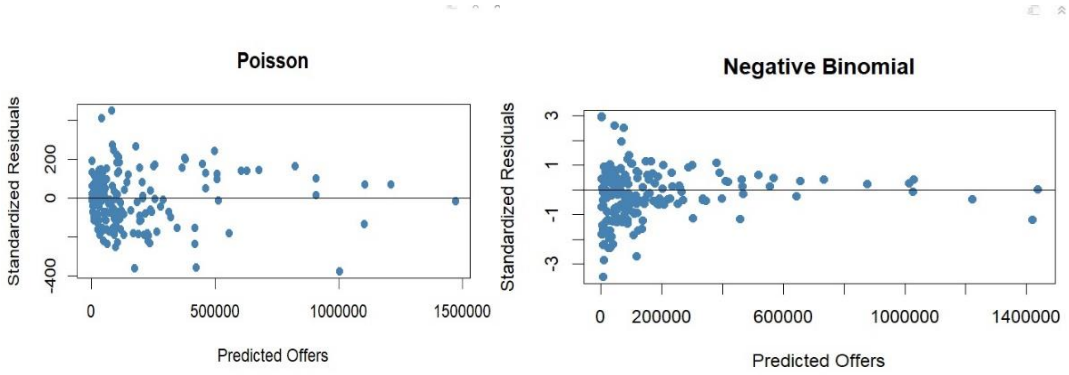

Figure 3: Standardized residual plot to compare the Poisson and Negative Binomial Models.

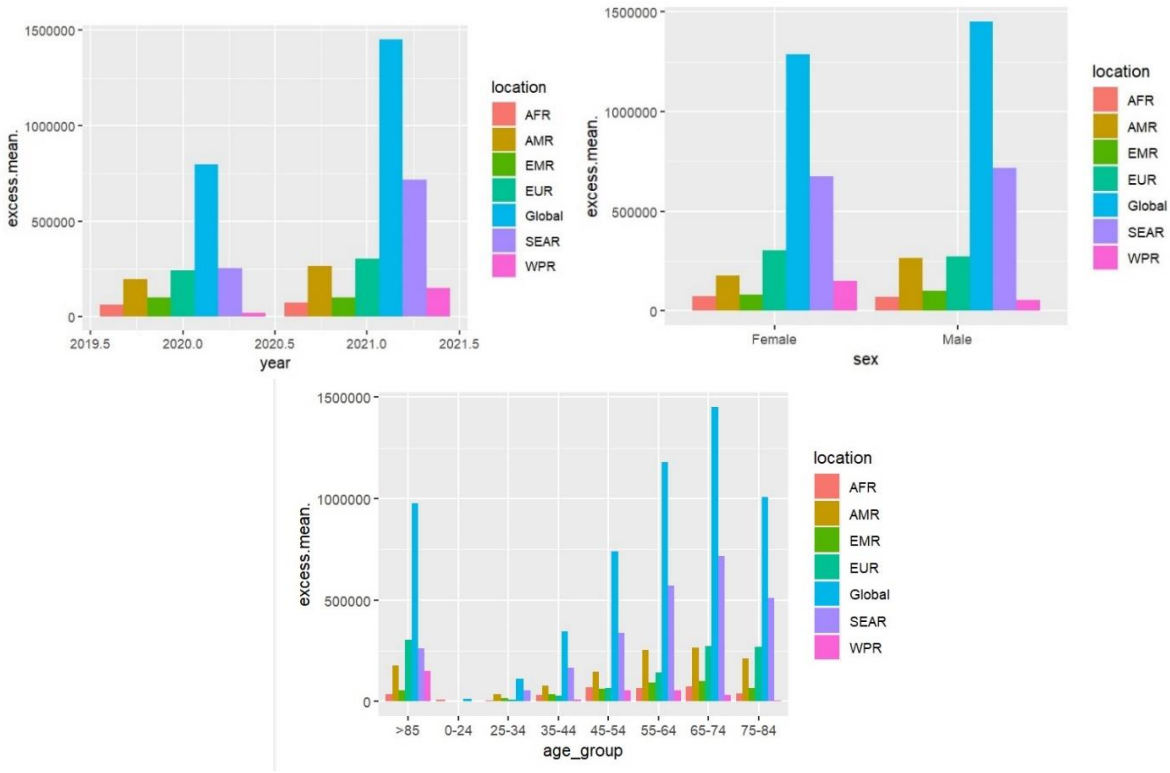

Figure 4: Histogram plot for each level of predictors

Table 7: Variance Inflation factor (VIF) (Negative Binomial Model)

locationAMR	locationEMR	locationEUR	locationGlobal
1.820896	1.774284	1.800057	2.991252
locationSEAR	locationWPR	year	age_group0-24
1.816763	1.553423	1.057914	2.640302
age_group25-34	age_group35-44	age_group45-54	age_group55-64
1.586736	1.852535	1.848639	1.788165
age_group65-74	age_group75-84	sexMale	pop
1.738003	1.676559	1.018429	3.250759

This is the result of the Variance Inflation factor (VIF), which is used to test multicollinearity (Raihen, N et al., 2017).

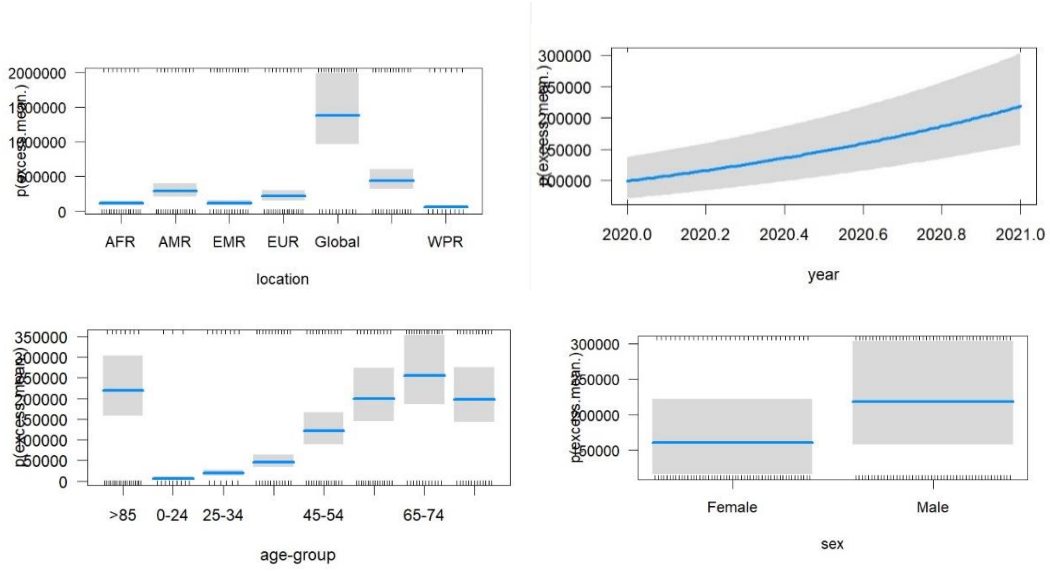

Figure 5: Plots of Comparing According to levels of explanatory variables.

The figures mentioned above illustrate that the levels of each factor correspond to which belongs to the largest and least positive excess mortality.

Table 7: An overview of the statistically significant model (Negative Binomial)

Coefficient	P-value	95% CI	Comment
locationAMR	3.5e-09***	5.991170e-01 ,1.282005e+00 (does not include 0)	Statistically significant
locationEMR	0.833637	-3.001995e-01,3.630954e-01 (does include 0)	not statistically significant
locationEUR	2.7e-05***	3.143293e-01,9.962463e-01 (does not include 0)	Statistically significant
locationGlobal	< 2e-16***	2.038824, 2.967695 (does not include 0)	Statistically significant
locationSEAR	< 2e-16***	1.015143, 1.694157 (does not include 0)	Statistically significant
locationWPR	.000546***	-1.047437e+00,-2.47452e-01 (does not include 0)	Statistically significant
Year2021	< 2e-16***	6.112138e-01,9.69777e-01 (does not include 0)	Statistically significant
age_ group0-24	< 2e-16***	-4.121276e,-2.859378 (does not include 0)	Statistically significant
age_ group25-34	< 2e-16***	-2.870172,-2.026635	Statistically significant

		(does not include 0)	
age_ group35-44	< 2e-16***	-1.907238 , -1.206788 (does not include 0)	Statistically significant
age_ group45-54	.000287***	-9.29607e-01,-2.57504e-01 (does not include 0)	Statistically significant
age_ group55-64	0.546120	-4.22417e-01, 2.27447e-01 (does include 0)	Not Statistically significant
age_ group65-74	0.319546	-1.59962e-01, 4.749895e-01 (does include 0)	Not Statistically significant
age_ group75-84	0.539202	-4.17827e-01, 2.21539e-01 (does include 0)	Not statistically significant
sexMale	0.00030***	1.378433e-01, 4.826154e-01 (does not include 0)	Statistically significant
pop	0.049457~0.05	-1.596845e-09,2.418396e-10 (does include 0)	Not Statistically significant

It represents our statistically significant factors to estimate the positive excess mortality at COVID-19 at all causes of the effect.

4. Results and Discussion

Our research was to forecast the outcome using only the four input factors and the single output variable. In order to determine something, we carried out a series of models. Comparing the observed number of deaths with the mean number of deaths from our data, there were 47.28% female population and 52.71% male population. There were 14.13% African, 15.21 AMR,15.21% EMR, 15.76% EUR, 15.76% SEAR, 8.69% and Global 15.21%, there were 44.02% population from 2020 year, and 55.97% population from 2021, there were 47.28% female and 52.71% male, there was 14.67% population age > 85, 6.52% population age between 0-24, 7.60% population age between 25-34, 13.58% population age between 35-44, 14.67% population age between 45-54, 14.67% population age between 55-64, 14.67% population age between 65-74, 13.58% population age between 75-84, and 14.67% population age > 85 o to analysis positive excess mortality.

Table 1 shows an overview of the analysis data for the positive excess mortality. We found that there are no missing observations. **Table 2** summarizes the Poisson regression model. The coefficient result of location provides that significant mean positive excess mortality significantly increases under location factors AMR, EMR, EUR, Global, and SEAR compared to the factor of AFR when all other variables are held constant, and the log of mean of positive excess mortality significantly decreases under location factors WPR compared to the factor of AFR when all other variables are held constant, since (p-value < 0.05) for all factors of location has a significance level of predictor for positive excess mortality mean. For the year, the log of the mean of the number of positive excess mortality is higher, 0.777 under the year 2021 compared to the year 2020 when all other variables are held constant, and since p-value <0, so the year has a significant factor for the excess mortality of COVID-19.

The coefficient of sex shows that the log of the mean of the male positive excess mortality mean increased by 0.23 times compared to the female when all other is held constant, and through a p-value<0.05, sex has a significant effect on the excess death at COVID-19. The coefficient values of age-group show that for the age groups 0-24,25-34,35-44,45-54, the log of the mean decreases compared to the age group more than 85 age, and the age-group shows that for the age group 55-64,65-74,75-84, the log of mean of positive excess mortality increases compared to age group more than 85 age, and age is a significant factor for the excess mortality of the presence of COVID pandemic (p-value<0.05). Population has a positive effect and is a significant predictor for our analysis of the positive excess mortality mean. (Thus, the coefficient is $1.525*10^{\wedge -10}$, and the p-value<0.05). **Table 4** checking the overdispersion of our Poisson regression model, we found the p-value< 0.05, which means that overdispersion exists, which violates the assumption that the mean and variance are the same in Poisson regression.

In the presence of overdispersion, we applied the Negative Binomial regression model as an alternative to the Poisson regression model for our positive excess mortality data at COVID-19. **Table 3** is the summary of the negative binomial regression model result. Results of "R" are displayed in **Table 3;** we find the regression coefficients for each of the variables along with the standard errors, z scores, and P-values. The indicator variable of location shows that the expected difference in log count of positive excess mortality under the locationAMR, locationEMR, locationEUR, locationGlobal, locationSEAR higher than the locationAFR(as a reference group), and the expected difference in log count of positive excess mortality under locationWPR is lower than the locationAFR group when all others variables are held constant, whatever p-values for all indicators do not means they are significant.

The variable year has a coefficient of 0.790, which is statistically significant; hence p-value<0.05, and it means that for each one-year increase in the year, the expected log count of the number of positive excess death increases by 0.0790 when held all other predictors are the same. The coefficient of sex from the Negative Binomial mode provides that the expected difference in log count of positive excess mortality considering the male population is 0.3104 higher than the female gender when all other predictors remain the same. The factorial variable of the age-group represents that for all age groups except the age -group 65-74 that the expected difference in log count of positive excess mortality is lower than the age-group of more than 85 (as a reference group) when keeping all other predictors are the same. And the coefficient of pop is $-7.30*10^{-10}$ (p-value=0.05), which means that for each one unit increase in population, the expected log count of the number of positive excess death decreases by $7.30*10^{-10}$ while taking all other predictors are the same, and it is not a statistically significant predictor for excess mortality at COVID-19.

From **Table 6**, thus the AIC and residual deviance of the Negative binomial model are both much smaller than those of the Poisson regression model, so we can choose the Negative binomial model as a better fit model for our data of COVID-19 for estimating the positive excess mortality. **Figure 5** is the result of the Variance Inflation factor (VIF) to check the multicollinearity among the dependent variables; hence it shows that almost all predictors have VIF<10, which means that multicollinearity does not exist at the Negative Binomial model in this analysis, so we do not need standardized at a negative binomial model for our analysis.

5. Conclusion

Computed the additional mortality during the pandemic in 2020 and 2021 for the Global and WHO Member State Regions; Rather than providing new estimates, our goal was to critically fit more effective approaches and identify the essential elements influencing COVID-19's positive excess mortality, comparison, and limits, as well as the data set's implications (Raihen, M. N. I. et al., 2022). One of the strategies was used to display the results of this analysis (Shi Y. et al., 2020). AIC and residual deviance values demonstrate that Negative Binomial regression is a better model than Poisson regression for achieving the study's purpose. LocationAMR, locationEUR, locationGlobal, locationSEAR, locationWPR, year, age-group (0-24, 25-34, 35-44, 45-54), and sex of positive excess mortality are the most effective predictors pop has a P-value of 0.049457~ 0.05 and a 95% CI of (-1.596845e^-09, 2.418396e^-10), so we can consider it an insignificant predictor in this analysis. It also revealed that the Global of location predictors, the year 2021, age-group 65-74 has more positive excess mortality, and male has more positive excess mortality than female. In this regression analysis, the number of reported deaths from all causes during the COVID-19 pandemic was greater than the number of expected deaths, and location, age, time, and gender are more effective factors of the positive access mortality at COVID-19 in 2020-2021.

6. Study Limitations

Excess mortality data is unfortunately not available for many countries, and this is going to continue because of a lack of previous information. This is a big drawback when monitoring a global pandemic. Calculating excess mortality requires accurate, high-frequency mortality data from previous years. Typically, only wealthy nations that can afford high-quality data reporting systems have access to this data. More detailed data is needed to draw more specific and actionable conclusions about the positive excess mortality mean. Additional research could be taken in future to more precisely estimate the excess mortality from all causes attributable to the COVID-19 pandemic.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: https://orcid.org/0000-0003-2680-0658
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Aburto, J. M., Kashyap, R., Schöley, J., Angus, C., Ermisch, J., Mills, M. C., & Dowd, J. B. (2021). Estimating the burden of the COVID-19 pandemic on mortality, life expectancy and lifespan inequality in England and Wales: a population-level analysis. *J Epidemiol Community Health*, *75*(8), 735-740.

[2] Achilleos, S., Quattrocchi, A., Gabel, J., Heraclides, A., Kolokotroni, O., Constantinou, C., ... & Demetriou, C. A. (2022). Excess all-cause mortality and COVID-19-related mortality: a temporal analysis in 22 countries, from January until August 2020. *International journal of epidemiology*, *51*(1), 35-53.

[3] Basellini, U., Alburez-Gutierrez, D., Del Fava, E., Perrotta, D., Bonetti, M., Camarda, C. G., & Zagheni, E. (2021). Linking excess mortality to mobility data during the first wave of COVID-19 in England and Wales. *SSM-Population Health*, *14*, 100799.

[4] Bilinski, A., & Emanuel, E. J. (2020). COVID-19 and excess all-cause mortality in the US and 18 comparison countries. *Jama*, *324*(20), 2100-2102.

[5] Chambers, J., Hastie, T., & Pregibon, D. (1992). Statistical models in S. In *Compstat: proceedings in computational statistics, 9th symposium held at Dubrovnik, Yugoslavia, 1990* (pp. 317-321). Heidelberg: Physica-Verlag HD.

[6] Charro, F., Ali, A. H., Raihen, N., Torres, M., & Wang, P. (2023). A bifurcation phenomenon in a singularly perturbed two-phase free boundary problem of phase transition. *Nonlinear Analysis: Real World Applications*, *73*, 103911.

[7] Collaborators, C. E. M. (2022). COVID-19 Excess Mortality Collaborators: Estimating excess mortality due to the COVID-19 pandemic: A systematic analysis of COVID-19-related mortality, 2020-21. *Lancet*, *399*, 1513-1536.

[8] Faust, J. S., Krumholz, H. M., Du, C., Mayes, K. D., Lin, Z., Gilman, C., & Walensky, R. P. (2021). All-cause excess mortality and COVID-19–related mortality among US adults aged 25-44 years, March-July 2020. *Jama*, *325*(8), 785-787.

[9] Folino, A. F., Zorzi, A., Cernetti, C., Marchese, D., Pasquetto, G., Roncon, L., ... & Iliceto, S. (2020). Impact of COVID-19 epidemic on coronary care unit accesses for acute coronary syndrome in Veneto region, Italy. *American Heart Journal*, *226*, 26.

[10] Gobiņa, I., Avotiņš, A., Kojalo, U., Strēle, I., Pildava, S., Villeruša, A., & Briģis, Ģ. (2022). Excess mortality associated with the COVID-19 pandemic in Latvia: a population-level analysis of all-cause and noncommunicable disease deaths in 2020. *BMC public health*, *22*(1), 1-12.

[11] Ioannidis, J. P. (2021). Over-and under-estimation of COVID-19 deaths. *European journal of epidemiology*, *36*, 581-588.

[12] Islam, N., Shkolnikov, V. M., Acosta, R. J., Klimkin, I., Kawachi, I., Irizarry, R. A., ... & Lacey, B. (2021). Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries. *bmj*, *373*.

[13] Jha, P., Brown, P. E., & Ansumana, R. (2022). Counting the global COVID-19 dead. *The Lancet*, *399*(10339), 1937-1938.

[14] Karlinsky, A., & Kobak, D. (2021). Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *Elife*, *10*, e69336.

[15] Kiang, M. V., Irizarry, R. A., Buckee, C. O., & Balsari, S. (2020). Every body counts: measuring mortality from the COVID-19 pandemic. *Annals of internal medicine*, *173*(12), 1004-1007.

[16] Koffman, J., Gross, J., Etkind, S. N., & Selman, L. (2020). Uncertainty and COVID-19: how are we to respond?. *Journal of the Royal Society of Medicine*, *113*(6), 211-216.

[17] Krieger, N., Chen, J. T., & Waterman, P. D. (2020). Excess mortality in men and women in Massachusetts during the COVID-19 pandemic. *The Lancet*, *395*(10240), 1829.

[18] Mansfield, E. R., & Helms, B. P. (1982). Detecting multicollinearity. *The American Statistician*, *36*(3a), 158-160.

[19] Mungmunpuntipantip, R., & Wiwanitkit, V. (2021). The COVID-19 pandemic and traffic accidents. *South African Medical Journal*, *111*(3), 192-192.

[20] Németh, L., Jdanov, D. A., & Shkolnikov, V. M. (2021). An open-sourced, web-based application to analyze weekly excess mortality based on the Short-term Mortality Fluctuations data series. *PLoS One*, *16*(2), e0246663.

[21] Raihen, M. N., Akter, S., & Sardar, M. N. (2023). Women's Career Challenges and Opportunities (A Study of Career and Job Satisfaction among Bangladeshi Women). *Academic Journal of Research and Scientific Publishing*, *5*(51), 05-22.

[22] Raihen, M. N., Akter, S., & Sardar, M. N. (2023). Food Satisfaction among Students: A Study of Present Public University Students in Bangladesh. *Journal of Mathematics and Statistics Studies*, *4*(1), 01-18.

[23] Raihen, N. (2017). *Convergence Rates for Hestenes' Gram-Schmidt Conjugate Direction Method without Derivatives in Numerical Optimization* (Master Thesis, University of Toledo).

[24] Raihen, M. N. I. (2022). *A Bifurcation Phenomenon of Regularized Free Boundary Problems of Two-Phase Elliptic-Parabolic Type* (Doctoral dissertation, Wayne State University).

[25] Raihen, M. N., & Akter, S. (2023). Forecasting Breast Cancer: A Study of Classifying Patients' Post-Surgical Survival Rates with Breast Cancer. *Journal of Mathematics and Statistics Studies*, *4*(2), 70-78.

[26] Ramírez-Soto, M. C., & Ortega-Cáceres, G. (2022). Analysis of excess all-cause mortality and COVID-19 mortality in Peru: observational study. *Tropical Medicine and Infectious Disease*, *7*(3), 44.

[27] Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., ... & Roser, M. (2020). Coronavirus pandemic (COVID-19). *Our world in data*.

[28] Rocco, P., Rich, J. A., Klasa, K., Dubin, K. A., & Béland, D. (2021). Who counts where? COVID-19 surveillance in federal countries. *Journal of Health Politics, Policy and Law*, *46*(6), 959-987.

[29] Rozenfeld, M., Peleg, K., Givon, A., Bala, M., Shaked, G., Bahouth, H., & Bodas, M. (2021). COVID-19 changed the injury patterns of hospitalized patients. *Prehospital and disaster medicine*, *36*(3), 251-259.

[30] Salottolo, K., Caiafa, R., Mueller, J., Tanner, A., Carrick, M. M., Lieser, M., ... & Bar-Or, D. (2021). Multicenter study of US trauma centers examining the effect of the COVID-19 pandemic on injury causes, diagnoses and procedures. *Trauma Surgery & Acute Care Open*, *6*(1), e000655.

[31] Schöley, J. (2021). Robustness and bias of European excess death estimates in 2020 under varying model specifications. *MedRxiv*, 2021-06.

[32] Shi, Y., Wang, G., Cai, X. P., Deng, J. W., Zheng, L., Zhu, H. H., ... & Chen, Z. (2020). An overview of COVID-19. *Journal of Zhejiang University. Science. B*, *21*(5), 343.

[33] Stein Jr, I., & Raihen, M. N. (2023). Convergence Rates for Hestenes' Gram–Schmidt Conjugate Direction Method without Derivatives in Numerical Optimization. *AppliedMath*, *3*(2), 268-285.

[34] Stokes, A. C., Lundberg, D. J., Elo, I. T., Hempstead, K., Bor, J., & Preston, S. H. (2021). It is made available under a CC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder who has granted medRxiv a license to display the preprint in perpetuity. medRxiv preprint doi: https://doi.org/10.1101/2021.07.20.21260869; this version posted July 23, 2021. The copyright holder for this preprint 28 COVID-19 and excess mortality in the United States: A county-level analysis. PLOS Medicine, 18(5), e1003571.

[35] World Health Organization, https://www.who.int/data/sets/global-excess-deaths-associated-with-covid-19-modelled-estimates.

[36] World Health Organization, https://www.who.int/data/technical-advisory-group/covid-19--mortality-assessment.

[37] Yamaoka, K., Nakagawa, T., & Uno, T. (1978). Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations. *Journal of pharmacokinetics and biopharmaceutics*, *6*(2), 165-175.

[38] Zubiri, L., Rosovsky, R. P., Mooradian, M. J., Piper-Vallillo, A. J., Gainor, J. F., Sullivan, R. J., ... & Reynolds, K. L. (2021). Temporal trends in inpatient oncology census before and during the COVID-19 pandemic and rates of nosocomial COVID-19 among patients with cancer at a large academic center. *The Oncologist*, *26*(8), e1427-e1433.