RESEARCH PAPER

COVID-19 Risk Factor Identification based on Ohio Data

Qin Shao [1], [a] **Gerard Thompson** [a] **Amy Thompson** [b]

Coresponding author(s): [1] **qin.shao@utoledo.edu**

CC BY

[a] Department of Mathematics and Statistics, Toledo, Ohio 43606, USA, and [b] School of Population Health

In January COVID-19 was declared to be a global emergency and everyday life was disrupted. Many questions about COVID-19 remain to be answered. This paper provides an examination of the Ohio COVID- 19 data set. In particular, logistic regression is applied to the analysis of age and gender characteristics on the mortality of a patient. Based on the statistics and the p-values, gender and age play an important role in the outcome of a patient and the most vulnerable group is comprised of male patients who are more than eighty years old. This paper is an attempt to help in the formulation of public health policy towards confronting COVID-19 and paves the way towards a more comprehensive quantitative analysis as more data become available.

COVID-19 | logistic regression | odds ratio | mortality

Since December 2019 starting in China, COVID- 19 has been sweeping across the world bringing severe disruption to people's lives and the world economy. On January 31, the world health organization declared COVID-19 to be a global emergency. The principal means of transmission appears to be through the air and it seems to be more infectious than the annual wave of in influenza. As of July 31, the Johns Hopkins Pandemic website confirms 17,767,622 worldwide and 682,931 deaths for a mortality rate of 3.84%. In the United States 4,617,728 cases have been recorded with 154,320 deaths and mortality rate of 3.34%.

Researchers all over the world have been working on many aspects of COVID-19. These research interests range from investigating the biological mechanism of the virus for the purpose of prevention, treatment, and development of vaccine (4, 5, 6, 10, 16), to predicting the number of cases so as to make hospital bed and ventilator arrangements (2, 12, 13, 14, 15).

For example, as of July 31, the Centers for Disease Control and Prevention (CDC) cites 32 groups that were making predictions for the coming four weeks. Among these groups are some of the most prestigious institutions in the world including Columbia University, Johns Hopkins University, the London School of Hygiene and Tropical Medicine and MIT. Among these 32 groups 15 employ SIRs methods (Susceptible, Infectious, Recovered) and variations of it including SEIR to which is added Exposed, that is, for patients who, although carry the disease, are asymptomatic. SIRs method involve coupled systems of ordinary differential equations. Some approaches involve difference equations or dynamical systems. Of the 32 groups referred to by CDC, eight use mainly or purely statistical methods including time series and three use machine learning techniques. In (13), growth of the epidemic was modeled using Verhulst's growth differential equation, which is discussed in (2). Its solutions involve logistic curves that are typically S-shaped and serve as models that "flatten the curve".

However, if we consult Figure 1, we see that it is doubtful whether in Ohio, the curve has indeed yet been flattened. The method is combined with a non-linear least squares approach to estimate parameters and the results applied to the growth of COVID-19 in a number of countries. A similar approach is applied to study the development of the disease in China (12). The principal concern here also involves a logistical model, but one that comes from statistics, as we shall now explain.

In this paper we formulate a logistical model for the growth of COVID-19 based upon data gathered from the state of Ohio to identify risk factors. The Department of Health in Ohio, has been updating data for the COVID- 19 cases in the State of Ohio at the Coronavirus Dashboard (https://coronavirus.ohio.gov/wps/portal/gov/covid-19/dashboards). It contains information about cases and patients. It has been five months since the last death occurred on March 1, 2020 in Ohio. By July 31, 2020, there have

Submitted: 08/16/2020, published: 12/10/2020.

been 96,369 case counts and 3,665 death counts. In addition to the columns of the data set which is given in Table 1, the Dashboard provides a summary of the data, such as the county, case map and cumulative case plots. However, the Dashboard does not include any statistical analysis. One of the goals of the current work is to provide more information by using statistical analysis. In particular, logistic regression is used to analyze the Ohio COVID-19 data according to age and gender and provide some insight into the mortality of patients. Risk factors will be identified from the publicly available data from the State of Ohio using statistical inference. Also, we will summarize the information of the data set, such as the mortality rate for each group of gender and age; and we will implemented logistic regression to make statistical inference to identify some risk factors.

Table 1. Ohio COVID-19 Data Columns

	Type	**Values**
County	factor with 88 levels	Adams, Allen, *etc*
Sex	factor with 3 levels	female, male, unknown
Age, range	factor with 9 levels	0 - 19, 20 - 29, 30 - 39, 40 - 49, 50 - 50 60 - 69, 70 - 79, 80+, Unknown
Onset date	factor with 203 levels	1/10/2020, 1/11/2020 *etc*
Date of death	factor with 139 levels	3/1/2020 *etc*
Admission date	factor with 156 levels	1/14/2020, *etc*
Case count	integer	0, 1, 2, . . .
Death count	integer	0, 1, 2, . . .
Hospitalized, count	integer	0, 1, 2, . . .

Materials and Methods

A data file was extracted including through July 31, 2020 from the Ohio Coronavirus Dashboard. However, it is important to note that the State of Ohio is constantly updating the data, including revising previously posted totals. The numbers used in this study is what were publicly available as of August 10. Age, Sex, Case Count, Death Count are the variables which will be studied in this paper and which will sometimes also be referred to age, gender, case, mortality". The Coronavirus Dashboard provides very detailed definitions of these variables. For example, a patient was counted as a case if she/he was confirmed or met the CDC Expanded Case Definition (Probable); a death was counted if it was considered to be COVID-19 related. From Table 1, the factor gender has three levels (female, male, unknown), and age has nine levels (0 - 19, 20 - 29, 30 - 39, 40 - 49, 50 - 59, 60 - 69, 70 - 79, 80+, unknown). The entire data set was summarized, including 784 cases with either gender unknown or age unknown or both. From now on, we will simply exclude these 0.81% of the cases with missing information, and analyze the rest of the data.

Table 2 is the summary for the case counts of all of the age-by-gender groups. For example, the high- est case counts are respectively 10,763 for females and 8,776 for males in the age range of 20 - 29, whereas the lowest case count is 3,390 for females in the age range of 70 - 79 and 2,259 for males in the age range of 80+. A relative frequency (RF) for each age-by-gender group is calculated by:

$$RF = \frac{GroupCount}{TotalCount} \qquad [1]$$

The Count in [1] above can be either a case count in Table 2 or death count in Table 3. From figure 2, the case relative frequencies are distributed fairly evenly between 0.0490 and 0.2174 for both genders and eight age groups. However, the death counts in Table 3 show an obvious upward trend in age for both females and males. The death count increases from 2 in the age range of 0 - 19 to 1,118 in the age range of 80+ for the female, and from 0 to 790 for the male. If only age is taken into account, the relative frequencies and cumulative relative frequencies (CRF) for the eight age levels in Table 3 suggest that about 91% of all the deaths were of people 60 years old and older. The female death counts and male death counts share the same rising pattern, namely, the percentage among all the deaths increases as age becomes bigger. For example, more than 61% of all the female deaths and more than 42% of all the male deaths were of patients 80 years old and older. There are big jumps in the relative frequencies for females and males in the age range of 80+ in both genders, as one may see in figure 3. In particular, the jump is almost 40% for females and accounts for nearly 62% of the female deaths in the age range of 80+.

The mortality probabilities of all the age-by- gender groups are our primary concern and will be estimated using a logistic regression model in the next section. A mortality rate (MR) is a relative frequency which is defined as the ratio of death count to case count:

$$MR = \frac{DeathCount}{CaseCount} \qquad [2]$$

The mortality rates in Table 4 exhibit a pronounced upward trend, as shown in figure 4(a). The risk or the death likelihood becomes bigger for an older patient. The total of the two mortality rates of females and males in the same age range increases from 0:05% for the youngest age level, to 61:85% for the oldest age level.

The total mortality rates of the female and the male gender levels are broken down by age in figure 4(b), and the pink bars corresponding to 80+ years old are obviously largest for both genders.

Results

In this section, COVID-19 data is analyzed using a logistic regression model to provide comparisons between mortality probabilities of the age-by- gender groups. The response variable *y* will be 1 if the patient is dead and 0 otherwise. The variable *y* is binary, and the mortality probability π is defined as π = Prob(y = 1). We are interested in whether π depends on either age or gender or both. After deleting "unknown", there are eight levels for the factor age and two levels for the factor gender. Thus, there are a total of 16 different groups, and each case is classified into one of these groups based upon gender and age. We use π_{ij} (i = 0; 1; j = 0; . . . ; 7) to denote the mortality probability of a patient whose gender is *i* in the *j*th age group. In particular, we define *i* = 0 for the female level of gender and *i* = 1 for the male level; *j* = 0 for the age level of 0 - 19, *j* = 1 for the age level of 20 - 29, *j* = 2 for the age level of 30 - 39, and so on.

The eight levels of age are coded by seven in- dicators (age_2; age_3; . . . ; age_8) with age_i = 1 if the case was in the *j*th age level, and the two levels of gender are coded as an indicator $gender_M$ with $gender_M$ = 1 if the case is in the *i*th gender level. We do not need any indicators corresponding to the reference category, which is the 0 - 19 age level and female gender level, for reasons that will become clear. Using an indicator for each level is called dummy coding, which greatly simplify statistical inference and interpretations.

A logistic regression model, which is a type of generalized linear model, is commonly used to describe the relationship between a binary response variable and independent variables. Some examples of the vast applications of logistic regression models are, the extent to which maternal drinking affects baby birth defects (7); how the result of presidential elections is related to the gross domestic product and other economic indicators (9). The books (8) and (1) are very detailed references for both the theory and applications of generalized linear models. In the context of the Ohio COVID-19 data, using dummy coding for the factors age and gender, the logistic regression model is defined as follows:

$$logit(\pi_{ij}) = log(\frac{\pi_{ij}}{1-\pi_{ij}}) \quad \textbf{[3]}$$

$$= \beta_0 + \beta_1 gender_M + \beta_2 age_2 + ... + \beta_8 age_8 \quad \textbf{[4]}$$

where β_0; β_1; β_2; . . . ; β_8; are unknown parameters and will be estimated from the data. Unlike a linear regression model, a logistic model describes the relationship between the odds π = /(1 - π) and the independent variables (gender, age). Larger odds implies that it is more likely for the event, which is death for the COVID-19 data, to happen. According to the formulation of the model [2], every β_k (k = 1; . . ; 8) represents the difference of log odds of two groups.

For example, β_1 is the log odds difference or log odds ratio,

$$logit(\pi_{1j}) - logit(\pi_{0j})$$

between males and females in the same age range. It is straightforward that $e^{\beta\ i}$ indicates the odds ratio of two groups. A more detailed explanation of the implications of these parameters can be found in Table 5. The "Interpretation" column of Table 5 is the meaning or implication of the parameters from a mathematical derivation based on the setup of the model [2].

Dummy coding not only facilitates interpretation of the model parameters, but simplistic statistical inference. In particular, the mortality prob- ability difference of two age levels, boils down to whether or not $\beta_{\ j}$ = 0, j = 2; . . . ; 8, whereas the gender effect on the mortality probability is represented by 1. According to the p-values in Table 5, the effects of age levels are significantly different, except for the rst two age levels, 0 - 19 and 20 - 29 for which mortality probabilities are not statistically significantly different. For a xed age level, the difference of gender is also statistically significant.

In addition, the sign of the estimates for $\beta_{\ 1}$; . . . ; $\beta_{\ 8}$ implies the relationship of the mortality probabilities. For example, based on $\beta_{\ 1} > 0$, we have:

$$\beta_1 = \text{logit}(\pi_{1j}) - \text{logit}(\pi_{0j}) > 0$$

Equivalently,

$$\frac{\frac{\pi_{1j}}{1-\pi_{1j}}}{\frac{\pi_{0j}}{1-\pi_{0j}}} > 0 \quad \textbf{[5]}$$

Then we can conclude that:

$$\pi_{\ 1j} > \pi_{\ 0j}$$

which implies a male patient is more likely to die than a female patient. The estimates in Table 5 are not only positive, but are increasing for older age levels. We conclude that the MR is bigger for the male than the female, and the increase of MR becomes faster as age increases. In particular, a male COVID-19 patient in his eighties is 2927 times more likely to die than a female teenager patient. The increasing likelihood of death for an older patient is also shown in figure 5, where the estimated mortality probability in each category is calculated from a logistic regression model parameter estimate as follows:

$$\pi_{ij} = \frac{exp(\omega)}{1+exp(\omega)} \quad \textbf{[6]}$$

where:

$$\omega = \beta_0 + \beta_1 gender_M + \beta_2 age_2 + ...\beta_8 age_8$$

The mortality probabilities for male and female patients are about the same for the age ranges of 0 - 19 and 20 - 29, whereas the mortality probability of a male senior patient become significantly larger than a female senior patient. The Pearson's chi-squared test is often applied to a contingency table of two random variables to examine whether they are independent. It is equivalent to the test used in a logistic model for large sample sizes. In particular, it is calculated for the outcome y of a case and age as well as the contingency table of y and gender. Both p-values are much less than 0.05, which confirms that dependence of mortality on age and gender is statistically significant.

Discussion

In conclusion, one sees that there is a greater risk of mortality as age increases, with the greatest risk being in those over 80 years of age. There is also growing evidence to suggest that while equal numbers of men and women develop COVID-19, when looking across age groups, males are more likely to die with the exception of

the age group of 80+ where there are significantly more female than male deaths occurred. What is not accounted for in this analysis, is the difference in the numbers of females living to this age compared to males (7.87 vs 5.06 million) (Available at https://www.statista.com/statistics/241488/population-of-the-us-by-sex-and-age/). The results of this Ohio study are in accordance with observations from the national COVID-19 data that has been reported by age and gender (Available at https://www.statista.com/statistics/1127560/covid-19-incidence-rate-us-by-age-and-gender/).

The gender differences in COVID-19 deaths may also be linked to the higher percentage of pre-existing health conditions in males. In one study of 99 COVID-19 patients in China, the majority of these individuals were males and pre-existing health conditions such as COPD, diabetes, and heart disease (4). Moreover, there is also a gender difference in health risk behaviors such as males being more likely to use alcohol and tobacco. Finally, there are also underlying biological difference between men and women that make COVID-19 outcomes worse in men. Women in general have stronger immune systems than men and are better able to fend off infections. One study also found that estrogen was protective in female mice infected with a similar strain of the virus during the 2003 SARS outbreak (3). During that epidemic, men also had a much higher case fatality rate.

The results of this study are useful in predicting patient outcomes and helping to shape patient care policies or even the use of experimental therapies. The data that was analyzed for this work also could be combined with racial and ethnic data or even socioeconomic status for further examination. For example, in one study (11), whites were at a higher risk of COVID-19 due to the higher numbers of this population living to old age when compared to blacks. Moreover, in households where at least one worker was unable to work remotely, the risk of illness was increased. The option of working remotely is often linked to many white-collar jobs rather than those employed in lower paying and blue-collar jobs.

This study has several limitations that may restrict its applicability in other contexts. First, the data that was analyzed was a \snapshot" in time and not of a longer longitudinal nature. Second, there was no way to determine if those in certain age groups who died were in hard hit longterm care facilities, nursing homes, or correctional facilities. Third, due to the nature of the COVID-19 virus, the mortality rate may be under reported as many cases may not be confirmed at post-mortem. Lastly, there was no way to assess the extent of comorbidities such as hypertension or obesity that would increase the risk of death or act as a confounding variable.

Table 2. Summary of Case Counts

Age	**Gender**							**TC**	**RF**	**CRF**
	Female			Male						
	Count	**RF**	**CRF**	**Count**	**RF**	**CRF**	**Count**			
80+	4159	0.0840	0.0840	2259	0.0490	0.0490	71	6489	0.0673	0.0673
70-79	3390	0.0685	0.1525	3207	0.0696	0.1186	50	6647	0.0690	0.1363
60-69	5056	0.1021	0.2546	5549	0.1204	0.2390	73	10678	0.1108	0.2471
50-59	6890	0.1392	0.3938	7284	0.1581	0.3971	110	14284	0.1482	0.3953
40-49	6890	0.1392	0.5330	6964	0.1511	0.5482	106	13960	0.1449	0.5402
30-39	8008	0.1618	0.6948	7886	0.1711	0.7194	114	16008	0.1661	0.7063
20-29	10763	0.2174	0.9122	8776	0.1904	0.9098	148	19687	0.2043	0.9106
0-19	4328	0.0874	0.9996	4132	0.0897	0.9995	73	8533	0.0885	0.9991
U	21	0.0004	1.0000	23	0.0005	1.0000	39	83	0.0009	1.0000
TC		49505			46080		784			
RF		0.5137			0.4782		0.0081		Case Total = 96369	
CRF		0.5137			0.9919		1.0000			

U: Unknown; TC: Total Count; RF: Relative Frequency (RF = Cell Case Count / Total Marginal Case Count); CRF: Cumulative Relative Frequency (CRF = Total RF's).

Table 3. Summary of Case Counts

Age	Gender						TC	RF	CRF
	Female			Male					
	Count	**RF**	**CRF**	**Count**	**RF**	**CRF**			
80+	1118	0.6160	0.6160	790	0.4270	0.4270	1908	0.5206	0.5206
70-79	406	0.2237	0.8397	494	0.2670	0.6940	900	0.2456	0.7662
60-69	173	0.0953	0.9350	343	0.1854	0.8794	516	0.1408	0.9070
50-59	77	0.0424	0.9774	154	0.0833	0.9627	231	0.0630	0.9700
40-49	19	0.0105	0.9879	46	0.0249	0.9876	65	0.0178	0.9878
30-39	11	0.0061	0.9940	18	0.0097	0.9973	29	0.0079	0.9957
20-29	9	0.0050	0.9990	5	0.0027	1.000	14	0.0038	0.9995
0-19	2	0.0010	1.0000	0	0.0000	1.0000	2	0.0005	1.0000
TC		1815			1850				
RF		0.4952			0.5048			Death Total = 3665	
CRF		0.4952			1.0000				

TC: Total Count; RF: Relative Frequency (RF = Cell Case Count / Total Marginal Case Count); CRF: Cumulative Relative Frequency (CRF = Total RF's).

Table 4. Summary of Mortality Rates

Age	Gender						Total		
	Female			Male			**Case**	**Death**	**MR**
	Case	**Death**	**MR**	**Case**	**Death**	**MR**			
80+	4159	1118	0.2688	2259	790	0.3497	6418	1908	0.2973
70-79	3390	406	0.1198	3207	494	0.1540	6597	900	0.1364
60-69	5056	173	0.0342	5549	343	0.0618	10605	516	0.0487
50-59	6890	77	0.0112	7284	154	0.0211	14174	231	0.0163
40-49	6890	19	0.0028	6964	46	0.0066	13854	65	0.0047
30-39	8008	11	0.0014	7886	18	0.0023	15894	29	0.0018
20-29	10763	9	0.0008	8776	5	0.0006	19539	14	0.0007
0-19	4328	2	0.0005	4132	0	0.0000	8460	2	0.0002
Total	49484	1815	0.0367	46057	1850	0.0402	95541	3665	0.0387

Mortality Rate (MR) = Death Count / Case Count.

Table 5. Logistic Model Parameters and Estimates

Model Parameter	Interpretation	Estimate	p-value
β_0	$\log\{\pi_{00}/(1\text{-}\pi_{00})\}$	-8:577	0.000
β_1	$\log[\{\pi_{1j}/(1\text{-}\pi_{1j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	0.421	0.000
β_2	$\log[\{\pi_{2j}/(1\text{-}\pi_{2j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	1.126	0.136
β_3	$\log[\{\pi_{3j}/(1\text{-}\pi_{3j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	2.042	0.005
β_4	$\log[\{\pi_{4j}/(1\text{-}\pi_{4j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	2.987	0.000
β_5	$\log[\{\pi_{5j}/(1\text{-}\pi_{5j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	4.240	0.000
β_6	$\log[\{\pi_{6j}/(1\text{-}\pi_{6j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	5.364	0.000
β_7	$\log[\{\pi_{7j}/(1\text{-}\pi_{7j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	6.511	0.000
β_8	$\log[\{\pi_{8j}/(1\text{-}\pi_{8j})\}/\{\pi_{0j}/(1\text{-}\pi_{0j})\}]$	7.561	0.000

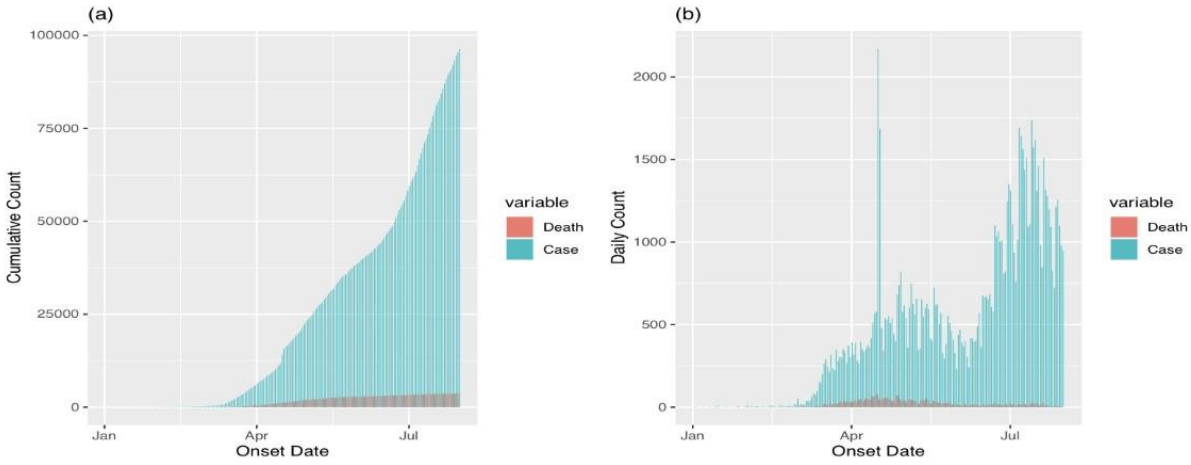

Fig. 1. Cumulative counts and daily counts

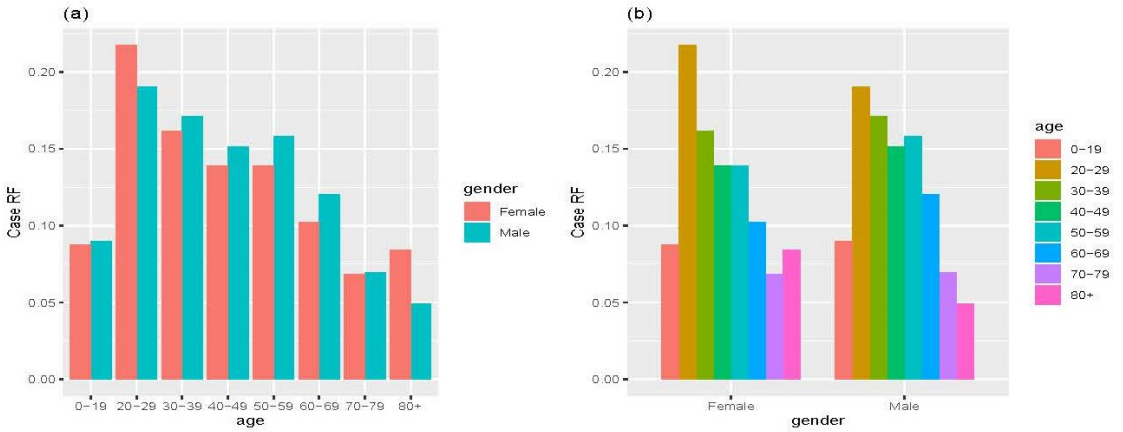

Fig. 2. Case relative frequency in each age by gender group

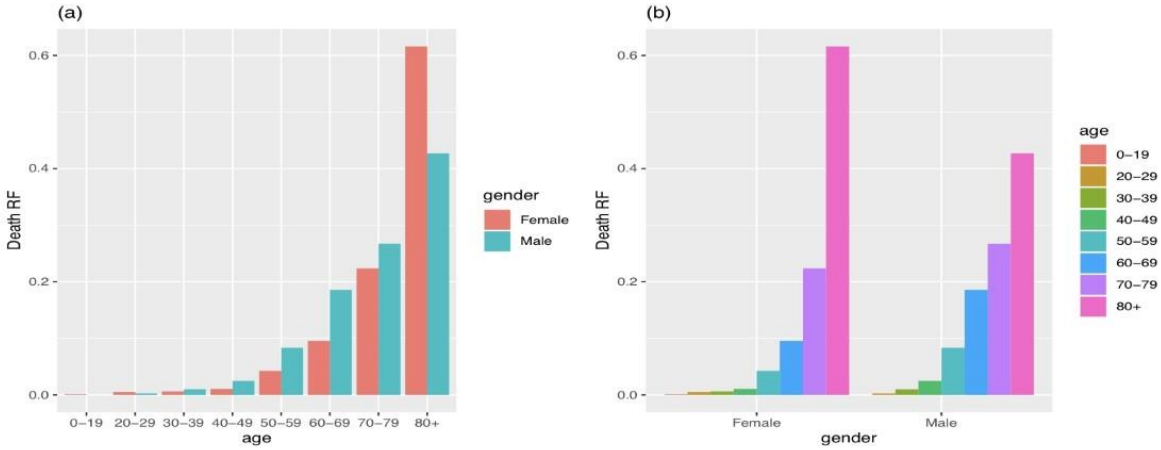

Fig. 3. Death relative frequency of each age by gender Group group

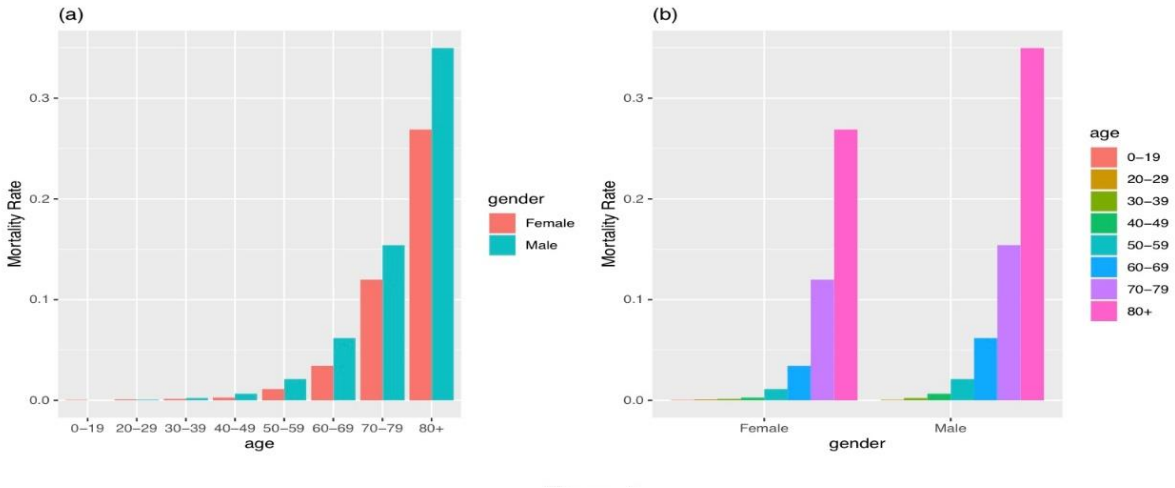

Fig. 4. Mortality rate of each age by gender group Group group

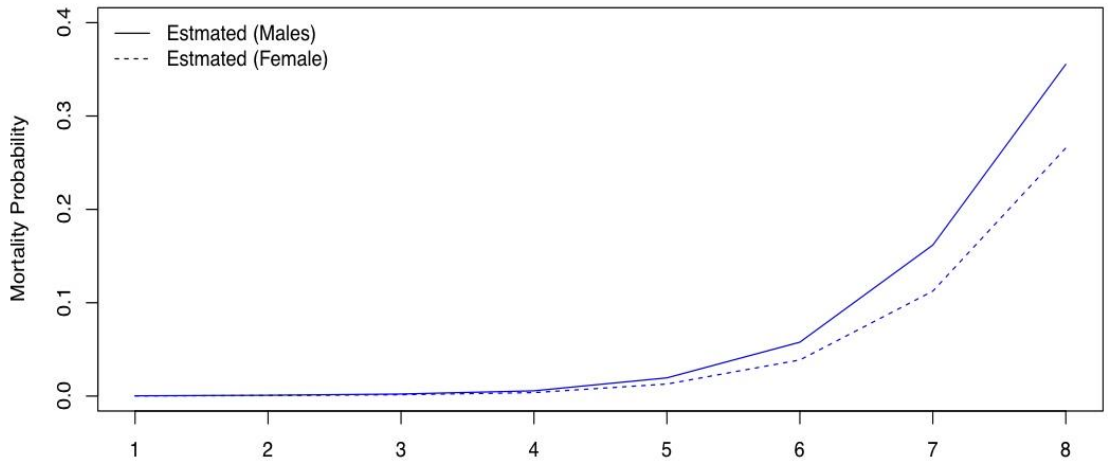

Fig. 5. Estimated mortality probabilities

Conclusion

In conclusion, one sees that there is a greater risk of mortality as age increases, with the greatest risk being in those over 80 years of age. There is also growing evidence to suggest that while equal numbers of men and women develop COVID-19, when looking across age groups, males are more likely to die with the exception of the age group of 80+ where there are significantly more female than male deaths occurred.

Conflict of interest

Authors declare no conflict of interest.

Authors' contributions

QS, GT proposed the research objective, QS performed data analysis, GT reviewed literature, AT provided the significance of the research from the public health perspective. All authors wrote the manuscript, read and approved the final document.

1. Agresti, A., Categorical Data Analysis (2nd), Wiley-Interscience, New Jersey, 2002.
2. Boyce, W. and diPrima R.C, Elementary Dif- ferential Equations (8th ed.), Wiley, 2005.
3. Channappanavar, R., Fett, C., Mack, M., Ten Eyck, P. P., Meyerholz, D. K., & Perlman, S. (2017) Sex-based differences in susceptibility to severe acute respiratory syndrome coronavirus infection. Journal of immunology, 198(10), 40464053.
4. Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y. et al. (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study, The Lancet, 395, 507-513. https://doi.org/10. 1016/S0140-6736(20)30211-7.
5. Chen, Y. Liu, Q., Guo, D. (2020) Emerging coronaviruses: Genome structure, replication, parthenogenesis, Journal of Virology, 92, 418- 423. doi:10.1002/jmv.25681
6. Guan, W. et al (2019) Clinical characteristics of coronavirus disease 2019 in China, The New England Journal of Medicine 382, 1708-1720. doi:10.1056/NEJMoa2002032
7. Graubard, B. I., and Korn, E. L. (1987) Choice of column scores for testing independence in ordered 2K contingency tables. Biometrics 43, 471-476.
8. McCullagh, P. and Nelder, J. A., Generalized Linear Models (2nd ed.), Chapman & Hall/CRC Monographs on Statistics & Applied Probability, 1989.
9. Nguyen, H. and Shao, Q. (2019) Logistic re- gression models with distributed lags, Journal of Data Science, forthcoming.
10. Paraskevis, D. et al (2020) Full-genome evo- lutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emer- gence as a result of a recent recombination event, Infection, Genetics and Evolution 79, 1- 4. doi:10.1016/j.meegid.2020.104212
11. Selden, T. M. and Berdahl, T. A. (2020) COVID-19 and racial/ethnic disparities in health risk, employment, and household composition, Health Affairs 39(9), 1-6. doi:10. 1377/hlthaff.2020.00897
12. Shen, C.Y., (2020) Logistic growth modelling of COVID-19 proliferation in China and its international implications, International Jour- nal of Infectious Diseases 96, 582-589. doi: 10.1016/j.ijid.2020.04.085
13. Wang, P., Zheng, X., Li, J. and Zhu, B. (2020) Prediction of epidemic trends in COVID-19 with logistic model and machine learning tech- nics, Chaos Solitons Fractals. 139: 110058, 1-7. doi:10.1016/j.chaos.2020.110058.
14. Xu, K. et al (2020) Application of ordinal lo- gistic regression analysis to identify the deter- minants of illness severity of COVID- 19 in China. Epidemiology and Infection 148, e146, 1-11. doi:10.1017/S0950268820001533.
15. Younes, A. B., Hasan, Z. (2020) COVID-19: Modeling, Prediction, and Control, Appl. Sci. 10, 3666. doi:10.3390/app10113666.
16. Yu, F., Du, L., Ojcius, D., Pan, C., Jiang, S. (2020) Measures for diagnosing and treat- ing infections by a novel coronavirus respon- sible for a pneumonia outbreak originating in Wuhan, China, Microbes and Infection 20, 74- 79. doi:10.1016/j.micinf.2020.01.003