Illness and Fatality Risks of COVID-19 of General Public in Hubei Provinces and Other Parts of China

Hongdou Li[1,3], Shuang Wang[2], Fan Zhong[1], Wuyin Bao[2], Yipeng Li[2], Lei Liu[1], Hongyan Wang[3], Yungang He[1*]

[1] Key Laboratory of Medical Epigenetics and Metabolism Institutes of Biomedical Sciences, State Key Laboratory of Genetic Engineering and School of Life Science, Fudan University, Shanghai 200032, China

[2] Guanghan Research Center of Personalized Healthcare, Shanghai 200438, China

[3]Obstetrics Gynecology Hospital, The Institute of Reproduction and Developmental Biology, Fudan University, Shanghai 200011, China

* Correspondence: heyungang@fudan.edu.cn

Abstract

New coronavirus 2019-nCOV poses a big challenge for global public health in early 2020. Coronavirus Disease 2019 (COVID-19) caused by the virus rapidly spreads all over the world and takes thousands of lives in just two months. To assess illness and fatality risk of the viral infection is exceedingly helpful to ensure effective management of the general public and patients in the outbreak. Therefore, it is critical to quantify illness and fatality risk of COVID-19 for the general public. In this report, we investigate the illness and fatality risk of the infection by analyzing the age composition of 5319 infected patients, 76 mortality cases, and 1,144,648 individuals of the general public in China. Our result shows a relatively low illness risk for young people but a

very high fatality risk for seniors. Notably, fatality risk could be as high as 0.48 for people older than 80 years. Furthermore, our study suggests that a good medical service can effectively reduce the mortality rate of the viral infection to 1% or less.

Introduction

On Jan 7, 2020, a new pathogenic virus caused pneumonia was identified in the sample of bronchoalveolar lavage fluid from a patient in Wuhan, Hubei province, China. It has typical features of the coronavirus family and therefore is classified as the subgenus Sarbecovirus, Orthocoronavirinae subfamily[1, 2]. Coronavirus is one of the main pathogens of human respiratory infection owing to frequent cross-species infections. This virus caused coronavirus disease in 2019 is the third epidemic coronavirus that emerges in the human population in the 21st century, following the severe acute respiratory syndrome coronavirus (SARS) outbreak in 2002 and the Middle East respiratory syndrome coronavirus (MERS) outbreak in 2012[3, 4].

The emerging virus rapidly becomes a challenge for global public health due to spread by human-to-human transmission. The majority of the earliest COVID-19 patients were linked to the Huanan Seafood Wholesale Market. However, human-to-human transmission has frequently been occurring and that the epidemic has been gradually growing[5]. As of February 22th, 2020, 76,396 laboratory-confirmed cases have been reported in China. Internationally, more than 1,515 cases have been reported in 28 countries[6, 7]. The number of infected individuals is far surpassing that of SARS and MERS.

This emerging virus can cause severe and even fatal respiratory diseases such as acute respiratory distress syndrome. It has been reported that COVID-19 is more likely to affect older males with comorbidities, suggests that age and comorbidity may be risk factors for poor outcomes[8, 9]. In China, the reported death is approaching 3% in total of COVID-19 patients in the middle of February 2019. As of February 21th, in China, 109 death cases are reported, while 399 new patients are identified. The death is still

increasing while the breakout is ending.

At present, information regarding the illness and case-fatality on clinical features and epidemiology of COVID-19 remains scarce. However, relatively accurate evaluation of illness risk and mortality is required that will help refine the risk assessment and ensure that the public and patients are managed in an effective way. Therefore, it is necessary to quantitatively evaluate the risks for individual groups of different ages and gender. In this report, we show our initial analysis of the public data from local authorities. Our study shows that the illness risk of COVID-19 might be as low as 0.1 for children while it could be over 0.9 for 40-years old adults. Our result also suggests that the fatality risk might be above 0.2 for patients older than 80 years old. Notably, the fatality risk is significantly different between patients of Hubei province and that of other parts of China.

Result & Discussion

Characteristic of identified patients and the general public

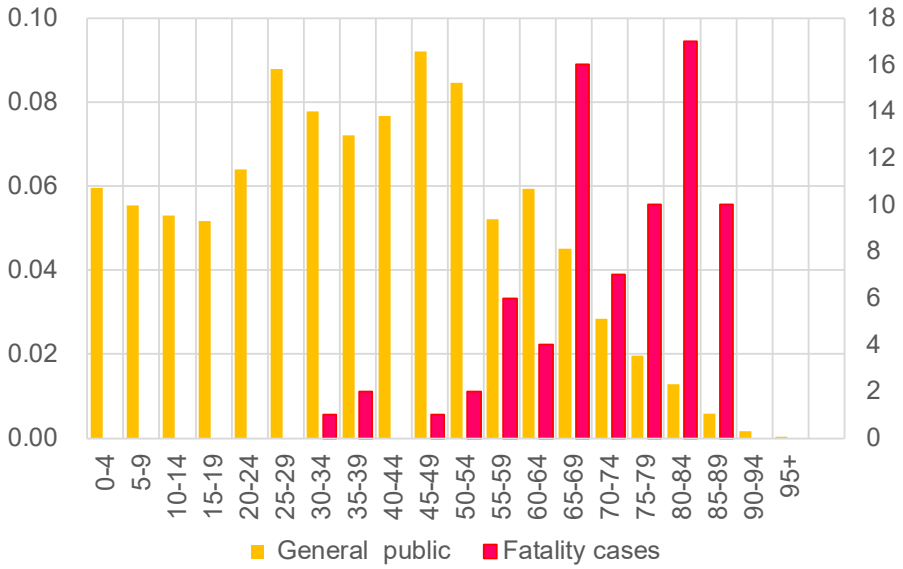
The public data of a total of 5319 identified COVID-19 cases were included in our analysis. There were 2829 (53.2%) males and 2490 (46.8%) females in the COVID-19 cases, the male to female ratio turned out roughly equal across all age groups. The age of COVID-19 patients ranged from 0.5 to 97 years, with a mean of 45.2 years. It should be noted that all the involved 5319 cases were reported by local authorities outside Hubei Province. Data of 1,144,648 individuals from the General Census of China (2018) was used as a composition of the general public[10]. The age and gender composition of COVID-19 patients and the public reference is presented in Table 1. Compared to the general public, the COVID-19 cases had higher average age, and there was a higher proportion of people aged 30 to 69 years.

Table 1 Age and gender composition of the general public and the identified COVID-19 cases.

	General public						COVID-19 cases					
Age Groups	**Total**	**Male**	**Female**	**Total (%)**	**Male (%)**	**Female (%)**	**Total**	**Male**	**Female**	**Total (%)**	**Male (%)**	**Female (%)**
0-4	67393	35887	31506	5.89	3.14	2.75	51	24	27	0.96	0.45	0.51
5-9	63322	34279	29043	5.53	2.99	2.54	59	35	24	1.11	0.66	0.45
10-14	62248	33775	28473	5.44	2.95	2 49	55	32	23	1.03	0.60	0.43
15-19	58258	31552	26706	5.09	2.76	2.33	95	55	40	1.79	1.03	0.75
20-24	68050	36085	31965	5.95	3.15	2.79	239	140	99	4.49	2.63	1.86
25-29	92977	47710	45268	8.12	4.17	3.95	356	204	152	6.69	3.84	2.86
30-34	93201	46843	46358	8.14	4.09	4.05	524	291	233	9.85	5.47	4.38
35-39	81886	41517	40370	7 15	3.63	3.53	567	305	262	10.66	5.73	4.93
40-44	83574	42557	41017	7 30	3.72	3.58	579	349	230	10.89	6.56	4.32
45-49	102384	52108	50276	8.94	4.55	4.39	662	354	308	12.45	6.66	5.79
50-54	96850	48939	47911	8.46	4.28	4.19	631	319	312	11.86	6.00	5.87
55-59	69844	35208	34636	6.1	3.08	3.03	494	240	254	9.29	4.51	4.78
60-64	68014	34092	33923	5.94	2.98	2.96	349	157	192	6.56	2.95	3.61
65-69	54799	26974	27825	4.79	2.36	2 43	311	147	164	5.85	2.76	3.08
70-74	34810	16905	17905	3.04	1 48	1.56	153	85	68	2.88	1.60	1.28
75-79	22799	10745	12054	1.99	0.94	1.05	96	46	50	1.80	0.86	0.94
80-84	14845	6457	8389	1.3	0.56	0.73	57	25	32	1.07	0.47	0.60
85-89	6902	2870	4033	0.6	0.25	0.35	28	14	14	0.53	0.26	0.26
90-94	2031	665	1365	0.18	0.06	0.12	10	7	3	0.19	0.13	0.06
95+	458	131	327	0.04	0.01	0.03	3	0	3	0.06	0.00	0.06
Total	**1144648**	**585299**	**559349**	**100**	**51.13**	**48.87**	**5319**	**2829**	**2490**	**100**	**53.19**	**46.81**

We collected detailed information of 76 mortality cases (45 cases from Hubei province and 31 cases from other parts of China) and plotted the age composition of the cases and the general public in Figure 1. It is evident that death occurs more frequently in older people but rare for patients under 40 years old. The mortality cases were from 34 to 89 years old, with an average age of 71.47 and a standard deviation 12.49. Due to the limit sample size, conditional effect of gender on ages was not investigated for the mortality cases in the present study.

Figure 1 Different age composition between the general public and 76 fatality cases. Individuals were grouped and presented on x-axis. For the general public, proportion of each age group was shown on left-hand side; number of fatality cases in each group was present on right-hand side.



Illness risk of COVID-19 for the general public

Based on the age composition of COVID-19 cases and the general public, we assessed the relationship between the illness risk and age in a Bayesian approach. Given the age composition of the general public and illness risk of different age groups, age

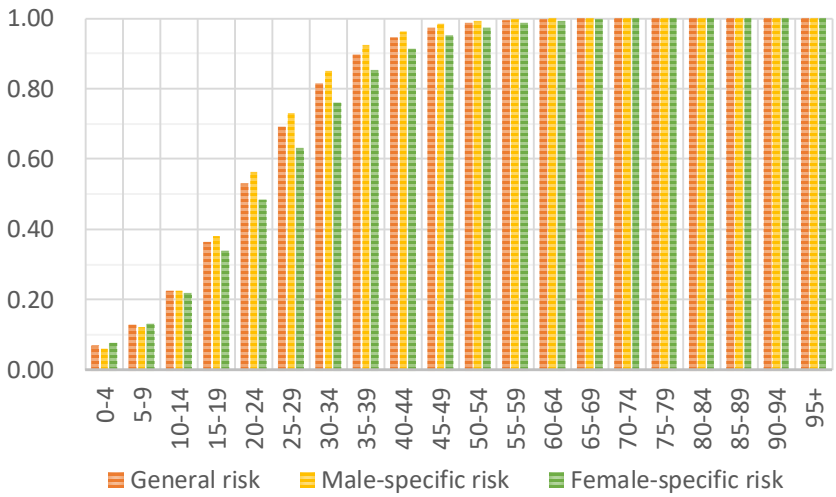
composition of infected individuals can be achieved in

$$\mathrm{P}(Age_i \mid \text{illness}) = \frac{\mathrm{P}(\text{illness} \mid Age_i) P(Age_i)}{\sum_i \mathrm{P}(\text{illness}, Age_i)}.$$

Where $\mathrm{P}(\text{illness} \mid Age_i)$ is the illness risk of age group *i*; $P(Age_i)$ is the proportion of age group *i* in general public. We assumed that the illness risk for different age groups could be obtained by a function of age, for example, a logistic function in this study. Therefore, the risk can be estimated in a maximum likelihood approach when the age composition of both COVID-19 cases and the general public is known.

Our result shows that the disease can happen in all age groups, and there lacks a significant difference between males and females (Figure 2). The illness risk is low for children and teenagers but rapidly increases for adults. For adults over 40-years old, the risk is higher than 0.9 when they have full exposure to the virus. The difference in illness risk for different gender is observed only for the groups between 15 and 50 years old. After the age of 15, males have slightly higher illness risk than women but the increase is ignorable for people over 50 years old. Our result does not support a previous report that COVID-19 generally affects more male than female in the epidemic[8].

Figure 2 Illness risk increases in older groups of the general public.



It should be mentioned that the illness risk in this study should be interpreted with the aforementioned Bayesian approach. Otherwise, it may mislead the understanding of our analysis. There is no evidence that any individual can be free from infection after full exposure to the virus. Infected young people may not become ill but asymptomatic during the course of infection.

Fatality risk of COVID-19 in general public

In a Bayesian approach similar to that mentioned above, we obtain

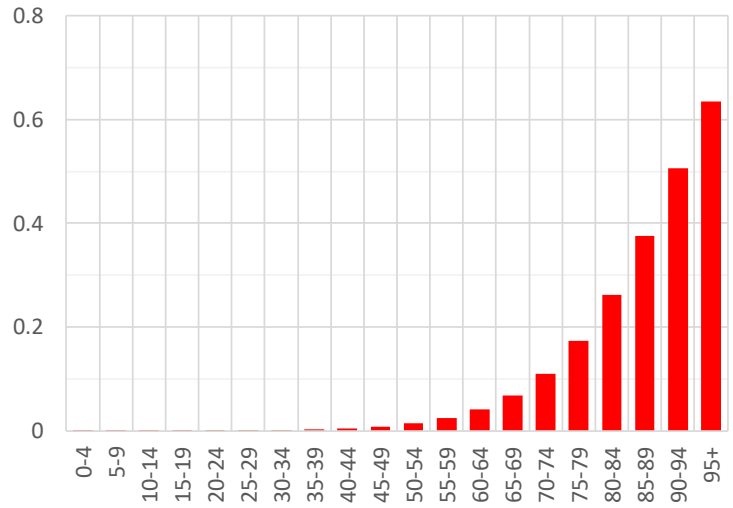
$$P(Age_i \mid Infection, Death) = \frac{P(Death \mid Infection, Age_i)P(Age_i \mid Infection)}{\sum_i P(Death, Age_i \mid Infection)}.$$

Where $P(Death \mid Infection, Age_i)$ is the probability of death condition on an individual's age and infection state; $P(Age_i \mid Infection, Death)$ is age composition of fatality cases of COVID-19. In this study, we assumed that infection happens in all age groups for the general public. In other words, we have $P(Age_i \mid Infection) = P(Age_i)$. With the assumption fatality risk of COVID-19 for the general public is a logistic function of age of each individual, we applied the maximum likelihood approach to estimate the fatality risk of COVID-19 in the general public $P(Death \mid Infection, Age_i)$.

Our result shows that there is a significantly higher fatality risk in older adults (Figure 3). The estimated fatal probability is less than 0.01 for individuals under 40 years old, but it is more than 0.51 for people older than 90 years. The estimated risk is much higher than the previous reports. To eliminate the concern that the high fatality risk of older people may inflate the mortality rate of infected population, we further imputed the mortality rate of infected people by summation of the proportion of death groups of different ages, i.e. $P(Death \mid Infection) = \sum_i P(Death, Age_i \mid Infection)$.Our result is consistent with most of the previous studies, supporting the hypothesis that older age is associated with an increased risk of mortality in COVID-19 patients. Age

has been reported as the independent predictor of adverse outcome in SARS and MERS. Comorbidities and low immune function of older people might be the major cause of higher mortality of coronaviruses[3, 4, 11]. Our analysis suggests a mortality rate of 2.38% for general infection. It agrees with previous reports for the raw mortality of COVID-19 in China. However, we noticed that raw mortality of COVID-19 is significantly different between identified cases of Hubei province and that of other parts of China.

Figure 3 Fatality risk dramatically increases for older age groups



Different fatality risk in Hubei and other provinces

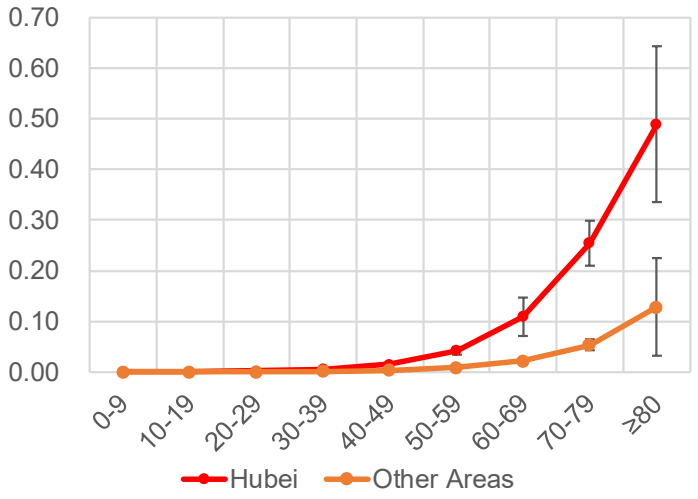
To compare fatality risk between Hubei and other provinces of China, we divided 76 mortality cases into two subsets, 45 cases from Hubei province and 31 cases from other parts of China. The aforementioned Bayesian analysis for fatality risk is applied to the two subsets with nine age groups each. We obtained standard deviation of estimations by applying the same method on 1000 simulated data sets that generated with initial estimation. Our result shows that fatality risk is no more than 0.13 ± 0.10 for people over 80 years old outside Hubei province, but the risk is as high as 0.60±0.15 for the corresponding age group in Hubei province. Fatality risk falls under 0.05 for people younger than 70 years in other parts of China, while only people under 50 years

old have a risk under 0.05 in Hubei province. We also calculated the mortality rate for a general infection inside and outside Hubei province as 4.78% and 0.95%. On the one hand, this difference may be partially explained by insufficient medical resources due to such a large amount of patients in Wuhan city in the breakout. On the other hand, detailed information on the majority part of fatal cases (40 of 45 in total) from Hubei province was published before January 25th, 2019. The fatality rate of early reported cases may be overstated because case detection is highly biased towards the more severe cases, and the risk may be decreased with the improvement of medical service.

Our analysis was based on the composition of the age of the different populations, and therefore it is less affected by the disease progression of patients, especially the increasing death of critically ill patients. As of February 23th 2020, there were still 10,968 COVID-19 patients in severe condition in China. In contrast to a row mortality rate of 2.3% reported by a research group of Chinses CDC, our analysis on their age data of mortality cases suggests a higher mortality rate as 5.63% for COVID-19 patients. It is higher than the fatality risk of infection for the general public in both Hubei province and other parts of China. The difference may be explained by the existence of asymptomatic carriers in infection cases of the general public.

In conclusion, we investigate the illness and fatality risk of the infection by analyzing the age composition of COVID-19 patients and the general public in China. Our data shows a relatively low illness risk for young people but a very high fatality risk for old adults. Therefore, it is prudent to strengthen the tertiary preventive and clinical care of old aged patients to reduce mortality. Furthermore, our results also support that a good medical service can effectively reduce the mortality rate of the viral infection to 1% or less. Our study could be of value to medical authorities to implement effective medical service. The lack of complete data for all COVID-19 cases potentially increases the occurrence of selection and measurement biases in this study. Therefore, further large-scale epidemiological studies are necessary to elucidate the risk factors of COVID-19 for the general public.

Figure 4 Estimated fatality risks in Hubei and other area of China. The risk was present on y-axis while age of grouped cases was shown on x-axis.



Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (Grant No. 31871255 and 91731310 to Y.H.) and Shanghai Municipal Science and Technology Major Project (Grant No. 2017SHZDZX01). The authors declare that there is no conflict of interest.

Reference

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, *et al.* China Novel Coronavirus Investigating and Research Team. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020.

2. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet %@ 0140-6736* 2020.

3. Donnelly CA, Ghani AC, Leung GM, Hedley AJ, Fraser C, Riley S, *et al.* Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *The Lancet* 2003, **361**(9371)**:** 1761-1766 %@ 0140-6736.

4. Ahmed AE. The predictors of 3-and 30-day mortality in 660 MERS-CoV patients. *BMC infectious diseases* 2017, **17**(1)**:** 615 %@ 1471-2334.

5. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y*, et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine* 2020.

6. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H*, et al.* First case of 2019 novel coronavirus in the United States. *New England Journal of Medicine %@ 0028-4793* 2020.

7. Rothe C, Schunk M, Sothmann P, Bretzel G, Froeschl G, Wallrauch C*, et al.* Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine %@ 0028-4793* 2020.

8. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y*, et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet %@ 0140-6736* 2020.

9. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J*, et al.* Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. *Jama* 2020.

10. National Bureau of statistics of China. *China Statistical Yearbook*, 2019.

11. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Chinese Journal of Epidemiology* 2020, **41**(2)**:** 145-151.