

# Modeling Survival Time to Recovery from COVID-19: A Case Study on Singapore

CURRENT STATUS: UNDER REVISION

BMC Public Health  BMC Series

Monira Mollazehi  
Hamad Medical Corporation

Mohammad Mollazehi  
Qatar University

Abdel-Salam G. Abdel-Salam  
Qatar University

✉ abdo@qu.edu.qa *Corresponding Author*  
ORCID: <https://orcid.org/0000-0003-4905-6489>

DOI:

10.21203/rs.3.rs-18600/v1

SUBJECT AREAS

*Infectious Diseases*

KEYWORDS

*Corona Virus, COVID-19, Survival Analysis, Parametric Model, Weibull Regression Model*

## Abstract

**Background:** Novel coronavirus (COVID-19) is a new strain of viruses that originated in China. In December 2019, a strange case of pneumonia was reported in Wuhan, which then was diagnosed to be COVID-19. In Singapore, the first positive case was reported on January 23, 2020.

**Aim:** This study aims to study the recovery time from COVID-19 in Singapore between January 23 and March 13, 2020.

**Methods:** It's a retrospective study from January 23 until March 13, 2020 for 187 cases with COVID-19 infection. Data of the instances were collected to identify the factors affecting the recovery time from COVID-19. Several parametric models were fitted and the best predictor model was selected using the Bayesian information criterion (BIC).

**Results:** Out of 187 patients, 96 (51.34%) were cured. The mean ( $\pm$ standard deviation) survival time was  $9.40 \pm 7.17$  days. Based on BIC, the exponential regression model was the weakest and the Weibull model was the best for fitting to data. According to the Weibull model, age (HR=1.01) and nationality (HR=0.76) had significant statistical effects ( $\alpha=0.05$ ) on patient recovery time.

**Conclusion:** According to the findings of the study, younger patients recover faster compared to elderly patients, and the Singaporean patients cure more quickly than non-Singaporean patients.

## Background

Corona Virus is a new human Betacoronavirus that uses densely glycosylated spike protein to penetrate host cells. The COVID-19 belongs to the same family classification with Nidovirales, viruses that use a nested set of mRNAs to replicate and it further falls under the subfamily of alpha, beta, gamma and delta Co-Vis. The virus that causes COVID-19

belongs to the Betacoronavirus 2B lineage and has a close relationship with SARS species. It is a novel virus since the monoclonal antibodies do not exhibit a high degree of binding to SARS-CoV-2. Replication of the viral RNA occurs when RNA polymerase binds and re-attaches to multiple locations (1,2).

Cases of COVID-19 started in December 2019 when a strange condition was reported in Wuhan, China. This virus has a global mortality rate of 3.4%, which makes it more severe in relation to flu. The elderly who have other pre-existing illnesses are succumbing more to the COVID-19. People with only mild symptoms recover within 3 to 7 days, while those with conditions such as pneumonia or severe diseases take weeks to recover. The recovery percentage of patients, for example, in China stands at 51%. The recovery percentage rate of COVID-19 is expected to hit 90% (3).

The virus has spread from China to 151 other countries and territories across the globe. From Wuhan, Hubei province, the virus spread to Mainland China, Thailand, Japan, South Korea, Vietnam, Singapore, Italy, Iran, and other countries (4). Singapore was one of the countries that were affected in the early stages of the COVID-19 spreading. The first positive case reported in Singapore was a Chinese national from Wuhan who tested positive for the virus on January 23, 2020 and the number continued to rise to reach up to 187 confirmed cases by March 13, 2020 (5).

This study aims to model the recovery time of patients infected with COVID-19 in Singapore, using several parametric models. Also, to examine the relationship between a patient's recovery time from COVID-19 and the contributing factors of age, gender, and nationality.

## Methodology

In this research, there are 187 cases tested positive for the novel coronavirus (COVID-19) between January 23, 2020 and March 13, 2020. These cases were registered in different

hospitals in Singapore. The data has been obtained from updates on COVID-19 by the Ministry of Health (6). The study used several parametric models, such as Exponential, Weibull, Logistic, Log-Logistic, Normal, Log-Normal, and Generalized Gamma were tested using SAS software. In order to compare those models and select the best-fitted model, the Bayesian information criterion (BIC) was utilized. The BIC is a measure of the goodness of fit. The model with a smaller value indicates that the proposed model fits the data well (7).

## Results

Out of the 187 patients with COVID-19 included in this study, 51.34% (96) of them were cured while the rest 48.66% (91) were censored. In this case, all censored instances considered to be right censored, which means they did not cure by the end of the study on March 13, 2020. The average time of a patient stay in the hospital was  $9.40(\pm 7.17)$  days. In addition, 57.52% (107) of patients were male, while 42.78% (80) were female. The average age of male patients was  $43.84 (\pm 16.56)$  years, while for female patients its was  $50.50(\pm 14.72)$  years.

**Table 1:** Demographic characteristics of the COVID-19 cases in Singapore

Characteristics		N (%)	Mean	SD
Age (year)	Male	107 (59.1)	43.84	16.56
	Female	80 (40.9)	50.50	14.72
	Total	187 (100)	46.69	16.10
Time (day)	Male	107 (59.1)	9.91	7.96
	Female	80 (40.9)	8.71	5.93
	Total	187 (100)	9.40	7.17
Age (year)	Singaporean	140 (74.87)	47.29	16.63
	Others	47 (25.13)	44.89	14.40
	Total	187 (100)	46.69	16.10
Time (day)	Singaporean	140 (74.87)	8.52	6.52
	Others	47 (25.13)	12.00	8.39
	Total	187 (100)	9.40	7.17

The BIC criterion for the exponential model was 749.43, which is the highest among all parametric models. For Logistic and Normal models, the BIC criterion was 736.78, and 735.47, respectively. Moreover, the BIC criterion for Log-Normal, Generalized Gamma, and Log-Logistic was 698.22, 697.54, and 695.53, respectively. Finally, the Weibull model had the lowest BIC with the value of 693.41, which infer that this model has a better fit to this data (Table 2). On the other hand, based on the plot of predicted probabilities against recovery time using Weibull distribution, it's clearly inferred that the Weibull regression model had a better fit to the data (Figure 1).

**Table 2:** Comparison of the BIC between Parametric Models

Model	Number of Parameters	BIC
Exponential	4	749.43
Weibull	5	693.41
Logistic	5	736.78
Log-Logistic	5	695.53
Normal	5	735.47
Log-Normal	5	698.22
Generalized Gamma	5	697.54

Based on the Weibull regression model, after adjusting other factors, the hazard ratio of age is 1.01 (95% CI: 1.01, 1.02), which is statistically significant at the level of 5%. Similarly, the hazard ratio of nationality was 0.76 and significant (95% CI: 0.61, 0.95). The only insignificant factor was the gender of patients (HR = 0.86, 95% CI: 0.61, 0.95).

**Table 3:** Diagnostic Factors of the COVID-19 Cases Using Weibull Regression Models

## Discussion

In modeling survival data, the semiparametric proportional hazards models are utilized frequently. In these models, no particular distributions are required for the data. Only an critical assumption called the PH assumption is needed to be checked (8). On the other

Parameters	HR*	Chi-Square	P-value	95% CI HR	
Intercept	8.93	209.58	0.00	6.69	12.06
Age	1.01	16.17	0.00	1.01	1.02
Gender					
Female	0.86	2.01	0.16	0.70	1.06
Male (Reference)					
Nationality					
Singaporean	0.76	5.62	0.02	0.61	0.95
Non-Singaporean (Reference)					

hand, parametric models might perform better. In this study, several parametric models were fitted and compared to each other using the BIC. Among seven parametric models, the Weibull model had the lowest BIC, which means that the Weibull model fitted this data better than the other parametric models

The objective of the study was to discover the factors that affect the recovery time of patients from COVID-19 infection in different hospitals in Singapore. Among the 187 patients that were included in this study, 51.34% were cured before March 13, 2020. Based on the results of the Weibull model, age and nationality are playing a significant role in survival time to recovery from COVID-19 infection. The results show that the hazard rate of recovery time from COVID-19 infection will increase by 1% as the age of patient increase by one year. In terms of nationality, the hazard rate of non-Singaporean patients was 0.76 times the Singaporean patients.

The main limitation attributed to the study is that it relies on a limited number of factors, which are gender, age, and nationality. Other demographic information and patients' vital signs were not obtainable due to the unavailability of these records online.

## Conclusion

In conclusion, the findings show that younger patients recover faster compared to elderly patients, and local patients cure more quickly than other patients. The study included all

the cases diagnosed with COVID-19 in Singapore between January 23, 2020 until March 13, 2020. Also, it is the first research that analyzed the factors affecting the recovery time from COVID-19 infection using several parametric models.

## Declarations

**Ethics approval and consent to participate:** Not applicable.

**Consent for publication:** Not applicable.

**Availability of data and materials:** Data is collected from

<https://www.moh.gov.sg/covid-19>

Also, the data can be collected from <https://www.gov.sg/article/covid-19-cases-in-singapore>

**Competing interests:** The authors declare that they have no competing interests.

**Funding:** Not applicable.

**Authors' contributions:** Authors participated in drafting the article, and they gave final approval of the version to be submitted.

**Acknowledgements:** Not applicable.

## References

1. McIntosh K. Coronavirus disease 2019 (COVID-19). UpToDate.  
<https://www.uptodate.com/contents/coronaviruses>. Accessed March 10 2020
2. Fisher D, Heymann D (2020) Q&A: The novel coronavirus outbreak causing COVID-19. BMC Medicine. doi: 10.1186/s12916-020-01533-w
3. World Health Organization. Coronavirus.  
<https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Accessed March 12 2020
4. Coronavirus Cases. <https://www.worldometers.info/coronavirus/#countries>. Accessed

March 14 2020

5. Singapore Government. Coronavirus disease 2019: Cases in Singapore.  
<https://www.gov.sg/article/covid-19-cases-in-singapore>. Accessed March 13 2020
6. Ministry of Health. Updates on COVID-19 (Coronavirus Disease 2019) Local Situation.  
<https://www.moh.gov.sg/covid-19>. Accessed March 11 2020
7. Weakliem DL (1999) A Critique of the Bayesian Information Criterion for Model Selection. *Sociological Methods & Research* 27:359–397
8. Kleinbaum DG, Klein M. *Survival Analysis: A Self-Learning Text*, 3rd ed. New York: Springer Science Business Media; 2013.

## Figures



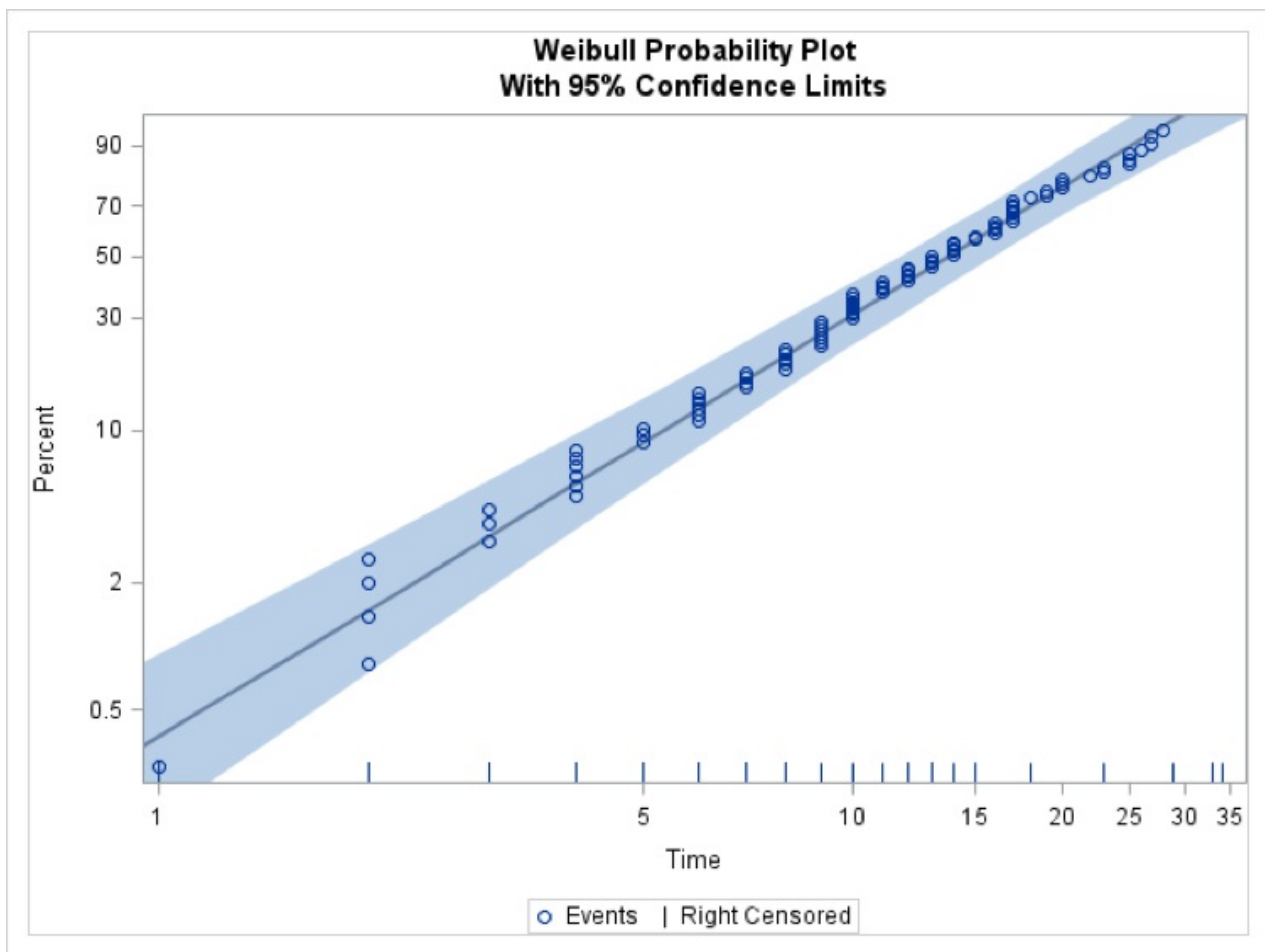


Figure 1

Weibull Probability Plot with 95% Confidence Limits for Recovery Time