Title	Brief Report: Incubation Period Duration and Severity of Clinical Disease Following Severe Acute Respiratory Syndrome Coronavirus Infection
Author(s)	Virlogeux, VML; Fang, J; Wu, JTK; Ho, LM; Peiris, JSM; Leung, GM; Cowling, BJ
Citation	Epidemiology, 2015, v. 26 n. 5, p. 666-669
Issued Date	2015
URL	http://hdl.handle.net/10722/220557
Rights	This is a non-final version of an article published in final form in Epidemiology, 2015, v. 26 n. 5, p. 666-669; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

BRIEF REPORT

Association of the incubation period duration with severity of clinical disease following Severe Acute Respiratory Syndrome coronavirus infection

Victor Virlogeux[1,2], Vicky J. Fang[2], Joseph T. Wu[2], Lai-Ming Ho[2], J. S. Malik Peiris[2,3], Gabriel M. Leung[2], Benjamin J. Cowling[2]

Affiliations

1. Department of Biology, Ecole Normale Supérieure de Lyon, Lyon, France.
2. WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.
3. Centre of Influenza Research, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.

Corresponding Author:

Dr Benjamin J Cowling, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Hong Kong Special Administrative Region, China

Tel: +852 3917 6711; Fax: +852 3520 1945; email: bcowling@hku.hk

Running head: Incubation period and severity

Word count (abstract): 138

Word count (main text): 1,472

FINANCIAL SUPPORT

This study was funded by the Harvard Center for Communicable Disease Dynamics from the National Institute of General Medical Sciences (grant no. U54 GM088558), and a commissioned grant from the Health and Medical Research Fund, Food and Health Bureau, Government of the Hong Kong Special Administrative Region. The funding bodies had no role in study design, data collection and analysis, preparation of the manuscript, or the decision to publish.

POTENTIAL CONFLICTS OF INTEREST

GML has received speaker honoraria from HSBC and CLSA. BJC has received research funding from MedImmune Inc. and Sanofi Pasteur, and consults for Crucell NV. The authors report no other potential conflicts of interest.

ACKNOWLEDGMENTS

The authors thank their colleagues in the Hong Kong Department of Health who were involved with the public health control of the SARS epidemic, and data collection and processing.

Association of the incubation period duration with severity of clinical disease following Severe Acute Respiratory Syndrome coronavirus infection

ABSTRACT

Background: Few previous studies have investigated the association between the severity of an infectious disease and the length of incubation period.

Methods: We estimated the association between the length of the incubation period and the severity of infection with the severe acute respiratory syndrome (SARS) coronavirus, using data from the epidemic in 2003 in Hong Kong.

Results: We estimated the incubation period of SARS based on a subset of patients with available data on exposure periods and a separate subset of patients in a putative common source outbreak, and we found significant associations between shorter incubation period and greater severity in both groups after adjusting for potential confounders.

Conclusions: Our findings suggest that patients with a shorter incubation period proceeded to have more severe disease. Further studies are needed to investigate potential biological mechanisms for this association.

INTRODUCTION

The incubation period of an infectious disease is the time from infection to onset of disease. Estimation of the incubation period of a novel pathogen can be vital for prevention and control, for example in order to determine appropriate duration of quarantine or observation of exposed persons.[1] In 2002-03, there was an epidemic of severe acute respiratory syndrome (SARS) caused by a novel coronavirus with more than 8,000 cases worldwide, mainly in Asia. The mean incubation period was rapidly estimated during the outbreak to be around 6.4 days,[2] and subsequent studies estimated a slightly shorter mean incubation period of around 4.0-5.3 days.[3–6] Estimation of the incubation period of a pathogen such as the SARS coronavirus can be complicated because infection events cannot be directly observed and exposure data are consequently interval-censored.[7]

The incubation period is thought to be a function of the initial infective dose, the speed of replication of the pathogen within the host, and within-host defense mechanisms.[1] Few previous studies have investigated the hypothesis that the incubation period might be correlated with the severity of disease, although some studies have examined the correlation between infecting dose and severity of disease.[8,9] Here, we analyze the association between the length of the incubation period and the severity of SARS using data from the 2003 outbreak in Hong Kong.

MATERIAL AND METHODS

Sources of Data

Information on all 1755 probable cases of SARS coronavirus infection were recorded in an electronic database extracted from a secure web-based data repository containing clinical and epidemiological data on all probable SARS cases admitted to hospitals in Hong Kong throughout the entire epidemic between February and July 2003.[10] Further details of the definition of a probable case of SARS and the database are reported elsewhere.[10,11] In a subset of cases, information was available on dates of exposure to infection, and in the majority of cases this was recorded as intervals of 2 or more days during which infection may have occurred rather than exact dates of infection.[2,6,7,10] We also analysed a separate subset of cases who were residents of the Amoy Gardens housing estate where a potential super-spreading event occurred in March 2003.[12–14] For these patients, we removed the small proportion of cases with onset dates prior to the main outbreak and with onset dates >14 days after the start of the main outbreak (**eAppendix**).

Statistical Analysis

A simple approach to estimate the incubation distribution from interval-censored data is to impute the midpoint of the exposure interval for each patient, and then estimate the distribution based on these 'exact' incubation times.[5,15] However, this approach is somewhat naïve, and is likely to overestimate incubation distributions which tend to be right-skewed.[3]

For the subset of cases with exposure dates, two approaches were used to estimate the incubation period distribution. First, we used a non-parametric estimator of the survival function that is a generalization of the Kaplan-Meier estimator for interval-censored data.[16] Second, we fitted a lognormal distribution[1,6,7,17] allowing for interval censoring, estimating the location and scale parameters (μ and σ) using Markov Chain Monte Carlo (MCMC) in a Bayesian framework.

To evaluate the association between the incubation period and the severity of disease, we first estimated the difference in mean incubation period between fatal and non-fatal cases in a Bayesian framework. However this analysis could not account for a potential confounding factor such as age which is known to be associated with the duration of the incubation period[5,6] and with the severity of disease.[10] We therefore specified a multivariable logistic regression model where death was the binary response variable and predictors included age, sex, occupation, and the incubation time T_i of each patient. We performed this with T_i resampled from the 10,000 posterior samples in each MCMC iteration. For the Amoy Gardens subset, we first estimated the potential date of infection for all cases by comparing the epidemic curve with the lognormal incubation period distribution estimated above, and then included in a similar logistic regression model.

In each of the analyses, we specified flat priors for each parameter, and drew 10,000 samples from the posterior distributions after a burn-in of 5,000 iterations. Further technical details of the statistical methods are provided in the **eAppendix**. All analyses

presented here were conducted using R version 3.1.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Among the 1755 probable cases of SARS in Hong Kong, 302 (17%) patients died.[10] The mean age was 44 years, the proportion of healthcare workers was 23% and cases that died were older and more likely to be male and not healthcare workers (Table 1). Among the 1755 cases, we identified 234 cases with an exposure period contained within the interval 0 to 20 days and 308 cases in the "Amoy Gardens" subset with an onset date within the interval 0 to 14 days (**eAppendix**). Both subsets had similar characteristics to the 1755 cases with fatal outcomes in 25/234 (11%) of the patients with exposure data and 38/308 (12%) in the Amoy Gardens subset (**Table 1**).

Parametric and nonparametric estimates of the incubation period distribution are presented in **Figure 1A** and show close agreement. We found a shorter incubation period for the fatal cases with a mean of 3.7 days (95% credibility interval, CrI: 2.6, 5.8), compared with a mean of 4.8 days (95% CrI: 4.2, 5.5) for the non-fatal cases, and a difference in means of 1.02 days (95% CrI: -0.41, 2.22) which was not significant.

The epidemic curve in the Amoy Gardens outbreak followed a very similar pattern, consistent with an infection event on 21st March 2003 (**Figure 1B**). Incubation periods for each patient were calculated based on this infection date. In this group, the mean incubation period was significantly shorter in the fatal cases 4.5 days (95% CrI: 3.8, 5.6) than in the

non-fatal cases 5.5 days (95% CrI: 5.2, 6.0) with mean difference 1.06 days (95% CrI: 0.16, 1.97) which was significant.

In the multivariable logistic regression model, we found that a shorter incubation period was generally associated with an increased risk of death in both subsets of patients. This association was statistically significant in the analysis of the patients with exposure intervals (OR=0.86; 95% CrI: 0.71, 1.00), and in the Amoy Gardens cluster with an OR=0.79 (95% CrI: 0.67, 0.94) (see also **eAppendix**).

To examine the sensitivity of our results to inclusion of patients with wide exposure intervals for the cases with exposure data, we also fitted the logistic regression models for a subset of 185 patients with shorter exposure intervals, and found very similar associations of the incubation period with risk of death (**eAppendix**). In addition, to examine the sensitivity of our results to the assumption of a linear association between incubation time and the log-odds of death, we categorized incubation times into tertiles and found similar results although the effects were only significant in the Amoy Gardens patients (**eAppendix**).

DISCUSSION

We estimated the incubation period of SARS based on two different subsets of patients, the latter with available data on exposure periods and the former with the hypothesis of a index patient contamination, and compared the length of this period between fatal cases and non-fatal cases to identify a correlation between shorter incubation and greater severity,

allowing for potential confounding by age, sex and occupation. Ours is the first study that examines the association between the incubation period and the severity of SARS in the literature to date. It is unlikely that a shorter incubation period itself is the cause of greater severity, but our results indicate that it could be a marker of underlying biological processes that led to greater severity. A shorter incubation period could be indicative of a higher infective dose, leading to faster/greater pathogen replication, out-running adaptive immune responses or leading to a more aggressive and damaging inflammatory response, and thus leading to more severe disease.

An association between severity and a shorter incubation period was suggested by Glynn et al.[8] in a study on malaria where a longer incubation period was associated with tertian fever, spontaneous recovery and less use of modifying treatment. Another study on salmonella infection reported a correlation between infecting dose and severity of the infection.[9] In a previous study, we showed that healthcare workers, who could have received a higher infecting dose, had a significantly shorter incubation period compared with non-healthcare workers.[6] It may also be biologically plausible that a more rapid progression from infection to symptom onset is correlated with more rapid disease progression after onset.

Our study does have some limitations, including a low number of fatal cases having available exposure data that may not be fully representative of all infections. In addition, exposures were self-reported and could be subject to recall bias which was differential and

affected by severity, and the cases in the Amoy Gardens cluster may not have all been infected on the same date.

REFERENCES

1. Nishiura H. Early efforts in modeling the incubation period of infectious diseases with an acute course of illness. *Emerg. Themes Epidemiol.* 2007;4:2.

2. Donnelly CA, Ghani AC, Leung GM, et al. Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet* 2003;361(9371):1761-1766.

3. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DAT. Incubation periods of acute respiratory viral infections: a systematic review. *Lancet Infect. Dis.* 2009;9(5):291-300.

4. McBryde ES, Gibson G, Pettitt AN, Zhang Y, Zhao B, McElwain DLS. Bayesian modelling of an epidemic of severe acute respiratory syndrome. *Bull. Math. Biol.* 2006;68(4):889-917.

5. Cai Q-C, Xu Q-F, Xu J-M, et al. Refined estimate of the incubation period of severe acute respiratory syndrome and related influencing factors. *Am. J. Epidemiol.* 2006;163(3):211-216.

6. Cowling BJ, Muller MP, Wong IOL, et al. Alternative methods of estimating an incubation distribution: examples from severe acute respiratory syndrome. *Epidemiology* 2007;18(2):253-259.

7. Farewell VT, Herzberg AM, James KW, Ho LM, Leung GM. SARS incubation and quarantine times: when is an exposed individual known to be disease free? *Stat. Med.* 2005;24(22):3431-3445.

8. Glynn JR, Bradley DJ. Inoculum size, incubation period and severity of malaria. Analysis of data from malaria therapy records. *Parasitology* 1995;110(01):7.

9. Glynn JR, Bradley DJ. The relationship between infecting dose and severity of disease in reported outbreaks of salmonella infections. *Epidemiol. Infect.* 1992;109(03):371.

10. Leung GM, Hedley AJ, Ho L-M, et al. The epidemiology of severe acute respiratory syndrome in the 2003 Hong Kong epidemic: an analysis of all 1755 patients. *Ann. Intern. Med.* 2004;141(9):662-673.

11. World Health Organization, Case Definitions for Surveillance of Severe Acute Respiratory Syndrome (SARS). *WHO*. Available at: http://www.who.int/csr/sars/casedefinition/en/. Accessed October 30, 2014.

12. Li Y, Duan S, Yu ITS, Wong TW. Multi-zone modeling of probable SARS virus transmission by airflow between flats in Block E, Amoy Gardens. *Indoor Air* 2005;15(2):96-

13. Yu ITS, Li Y, Wong TW, et al. Evidence of airborne transmission of the severe acute respiratory syndrome virus. *N. Engl. J. Med.* 2004;350(17):1731-1739.

14. Yu IT-S, Qiu H, Tse LA, Wong TW. Severe acute respiratory syndrome beyond Amoy Gardens: completing the incomplete legacy. *Clin. Infect. Dis.* 2014;58(5):683-686.

15. Centers for Disease Control and Prevention (CDC). Severe acute respiratory syndrome--Singapore, 2003. *MMWR Morb. Mortal. Wkly. Rep.* 2003;52(18):405-411.

16. Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. *J Roy Stat. Soc Ser. B* 1976;38:290 –295.

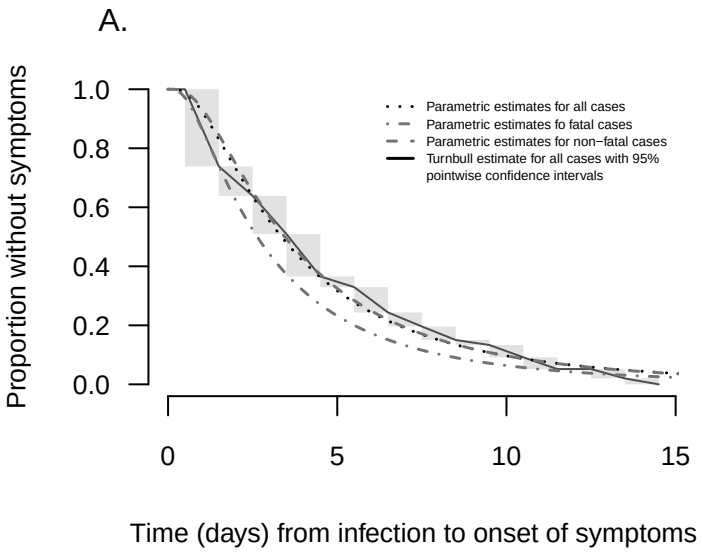
17. Sartwell PE. The distribution of incubation periods of infectious disease. *Am. J. Hyg.* 1950;51(3):310-318.

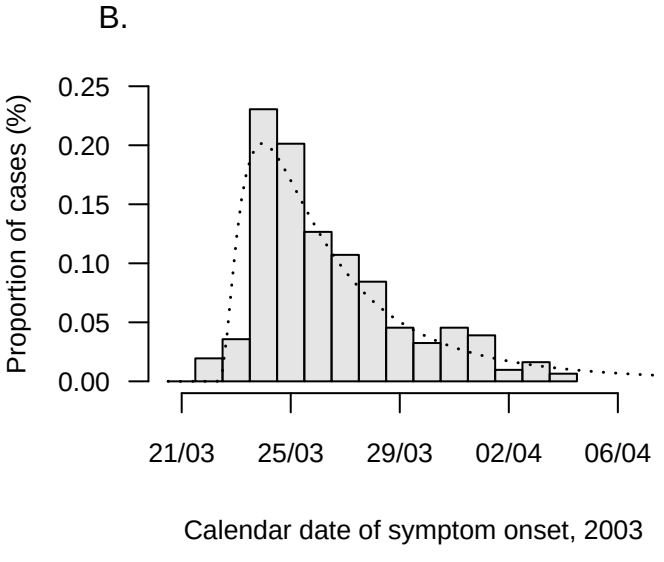
FIGURE LEGEND

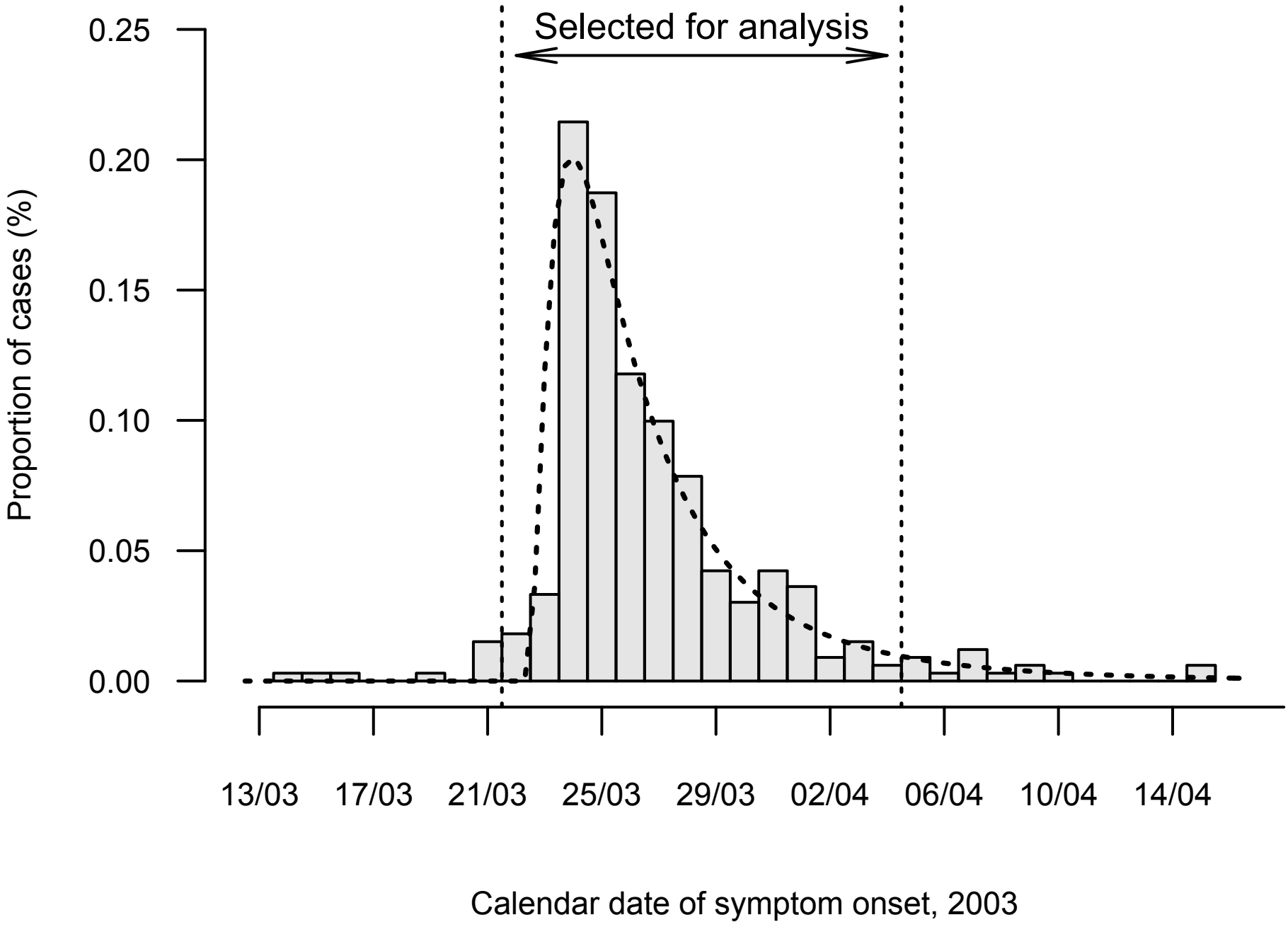
Figure 1. Panel (a): Parametric (dotted line) and non-parametric (solid line) estimates of the incubation distribution for SARS cases with available data on exposure times (n=234). The incubation distribution estimated with a lognormal model (dotted line) gives a mean incubation time of 4.7 days (95% credibility interval, CrI: 4.1-5.4 days) and a standard deviation of 4.6 days (95% CrI: 3.6-6.0 days) respectively. The non-parametric estimate of the incubation distribution is represented by a solid line, and gray rectangles show intervals where the nonparametric maximum likelihood estimate was not unique. **Panel (b): Distribution of illness onset dates for SARS cases in the Amoy Gardens cluster (n=308).** For this subset of patients, we hypothesized that all the cases were infected on 21 March 2003, and the epidemic curve is consistent with the lognormal incubation period distribution estimated on the separate subset of cases with exposure data as shown in panel A (i.e. mean 4.7 days, standard deviation 4.6 days).

Table 1. Characteristics of SARS patients

Patient characteristics	Fatal cases	Non-fatal cases	Overall	p-value
		All cases		
Sample size, n (%)	302 (17%)	1453 (83%)	1755	-
Age (years); mean±SD	66.6 ± 17.3	38.7 ± 17.3	43.5 ± 20.2	<0.001
Male sex, n (%)	173 (57%)	604 (42%)	777 (44%)	<0.001
Healthcare worker, n (%)	129 (43%)	276 (19%)	405 (23%)	<0.001
		Cases with exposure data		
Sample size, n (%)	25 (11%)	209 (89%)	234	-
Age (years); mean±SD	57.8 ± 14.7	40.1 ± 14.1	42.0 ± 15.2	<0.001
Male sex, n (%)	14 (56%)	99 (47%)	113 (48%)	0.546
Healthcare worker, n (%)	3 (12%)	54 (26%)	57 (24%)	0.202
		Amoy Gardens cases		
Sample size, n (%)	38 (12%)	270 (88%)	308	-
Age (years); mean±SD	48.4 ± 12.0	33.1± 14.1	35.0 ± 14.7	<0.001
Male sex, n (%)	22 (58%)	108 (40%)	130 (42%)	0.055


A.
Proportion without symptoms
1.0
0.8
0.6
0.4
0.2
0.0
Parametric estimates for all cases
Parametric estimates fo fatal cases
Parametric estimates for non−fatal cases
Turnbull estimate for all cases with 95% pointwise confidence intervals
0
5
10
15
Time (days) from infection to onset of symptoms


B.
Proportion of cases (%)
0.25
0.20
0.15
0.10
0.05
0.00
21/03
25/03
29/03
02/04
06/04
Calendar date of symptom onset, 2003


Selected for analysis
0.25
0.20
0.15
0.10
0.05
0.00
Proportion of cases (%)
13/03
17/03
21/03
25/03
29/03
02/04
06/04
10/04
14/04
Calendar date of symptom onset, 2003

eAppendix

Association of the incubation period duration with severity of clinical disease following Severe Acute Respiratory Syndrome coronavirus infection

Additional details of statistical methods

The incubation period is defined as the delay between infection and illness onset. If infection occurred at time X_i for the patient *i*, and symptom onset occurred at time Z_i, the incubation period is defined as $T_i = Z_i - X_i$. However, estimation of the incubation period is often complicated because infection events cannot be directly observed. If patient *i,* reported that infection most likely occurred in a period of exposure between times L_i and U_i, where $L_i \leq X_i \leq U_i$, the incubation time therefore is bounded by the interval ($Z-U_i$, $Z-L_i$). These data are a special type of survival data, and a natural approach would be to "reverse" the time axis setting Z as the origin and X as the outcome time. "Reversing" the time axis is valid only when the density function for infection is uniform in chronologic time. This condition should be reasonable here in the setting of SARS, with each exposure interval being relatively short.

Based on previous studies which identified the lognormal distribution as a satisfactory parametric model for the incubation period of SARS, in this study we assumed that the incubation distribution followed a lognormal distribution with parameters (μ, σ) and probability density function:

$$\boldsymbol{f(t_i) = \exp\left(-\left[\frac{\ln(t_i) - \mu}{\sigma}\right]^2 / 2\right) \Big/ \left(t_i \sigma \sqrt{2\pi}\right)}$$

We used a Bayesian framework to estimate the parameters of the incubation period distribution. In this framework, if θ represents a vector of parameters and y the data, Bayes theorem gives us the following relationship:

$$\boldsymbol{p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}} \quad \textbf{(1)}$$

where $p(\theta)$ is the prior probability of the parameters θ , $p(y|\theta)$ is the likelihood function and $p(\theta|y)$ is the posterior probability of θ given the data y.

The MCMC process was initiated by giving random values to the parameters θ and by choosing non-informative prior (flat prior) for θ. A Metropolis Hastings algorithm was used to update the parameter values in each iteration. In each iteration, all the k parameters are randomly generated using the normal distribution with the mean θ_k^{j-1} (previous value of the kth parameter) and standard error σ_k , $N(\theta_k^{j-1}, \sigma_k)$ for each parameter. The updated likelihood is compared with the previous one using the following accept-reject method:

$$\boldsymbol{q = \frac{p(y|\theta^j)p(\theta^j)}{p(y|\theta^{j-1})p(\theta^{j-1})}}$$

If $q \geq 1$, the proposed new values of parameters θ^j are accepted

If $q < 1$, then θ^j are accepted with probability q.

A burn-in period with 5,000 iterations was used to reduce the bias of the choice of the initial parameter values and to generate values only in the stationary distribution. The above algorithm was repeated 10,000 times after the burn-in period, with an acceptance rate included in [0.45,0.55] for each parameter (adjusting on σ_k).

For interval-censored data the following $p(y|\theta)$ was used:

$$\boldsymbol{p(y|\theta) = \prod_{i=1}^{n} S(u_i|\theta) - S(l_i|\theta)}$$

where $[u_i, l_i]$ are the interval censored data of the patient *i* and S(.) is the survival lognormal distribution function.

Parameters of the fitted lognormal distribution

In patients with exposure intervals data, the parameters of the lognormal distribution were estimated with the MCMC approach in the fatal and non-fatal cases, respectively. The estimates are for fatal cases: μ= 0.97 (95% CrI: 0.44, 1.38) and σ= 0.88 (95% CrI: 0.58, 1.34) and for non-fatal cases: μ= 1.24 (95% CrI: 1.09, 1.37) and σ= 0.88 (95% CrI: 0.58, 1.34).

Among patients of the Amoy garden cohort, the parameters of the lognormal distribution were estimated with the same approach approach. The estimates are for fatal cases: μ= 1.37 (95% CrI: 1.19, 1.54) and σ= 0.54 (95% CrI: 0.43, 0.68) and for non-fatal cases: μ= 1.59 (95% CrI: 1.53, 1.65) and σ= 0.51 (95% CrI: 0.46, 0.55).

Logistic regression model

The logistic regression model used in this study is based on the following equation:

$$ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_1(\textit{incubation period}) + \beta_2 X_2(age) + \beta_3 X_3(sex) + \beta_4 X_4(occupation)$$

where p_i is the probability of death, β_i's are the regression coefficients, estimated with MCMC (using flat priors N(0,100000)) and X_i 's the explanatory variables labeled directly in the equation above.

For the patients with exposure data, we performed this analysis three different ways: (1) with T_i as the imputed midpoint of exposure intervals; (2) with T_i as the mean of the 10,000 posterior samples i.e. simulated incubation times for each patient; (3) with T_i resampled from the 10,000 posterior samples in each MCMC iteration. Results of these different analyses are presented in **eAppendix Table 1**. Significant association between incubation period and risk of death was observed using the mean approach (2) and the imputation approach (3) whereas no significant association was observed with the midpoint

imputation method (1). The results of method (3) are presented in the main text, as this is the most appropriate approach.

Sensitivity analyses

Sensitivity analyses were conducted in both subsets of patients. For the patients with available exposure data, an interval of 0 to 20 days was selected for the main analysis due to the data reported by some patients with very wide exposure intervals that are not informative. As a sensitivity analysis, a subgroup of patients was extracted with an inclusion criteria based on the exposure intervals of no more than 10 days (N=185). We also fitted the logistic regression model for this subset of patients using the mean incubation time and found very similar associations of the incubation period with risk of death (OR=0.77; 95% CI: 0.53 - 1.04).

For the Amoy Gardens cluster of cases (**eAppendix Figure 1**), which were thought to have been caused by a super-spreading event, we examined potential dates of common infection for all members of the cluster and estimated that the most likely infection date was 21st March 2003. A period of 14 days i.e. cases with onset dates between 22 March 2003 and 4 April 2003 (shown in **eAppendix Figure 1**) was selected for the main analysis in order to exclude the small number of cases infected prior to the super-spreading event, and cases that had symptom onset late in the outbreak and may have been secondary or tertiary infections. We also conducted a sensitivity analysis using first a larger subset of patients, with onset dates between 22 March 2003 and 10 April 2003 (N=320) and secondly using a smaller subset with onset dates between 22 March 2003 and 31 April 2003 (N=286). Results from the logistic regression model show similar associations of the incubation period with risk of death (OR=0.92; 95% CI: 0.82-1.02 and OR=0.77; 95% CI: 0.62-0.94, respectively).

To evaluate the potential cofounding effect of age, we conducted a stratified analysis using two categories of age with the threshold of 50 years (**eAppendix Table 2**). A significant

association between incubation period and risk of death was only observed among Amoy garden cohort patients. No significant association was observed for patients with exposure data, probably due to the reduced sample size in the sub-analysis.

Moreover, to examine the sensitivity of our results to the assumption of a linear association between incubation time and the log-odds of death, we conducted a similar analysis with a logistic regression model described previously with the incubation period as a categorical variable, using tertiles (**eAppendix Table 3**). We observed a significant association between shorter incubation period and an increased risk of death only in the Amoy Garden cohort. However, no significant association was observed in the cases of the exposure data patients although the direction of point estimates was consistent with the analysis presented in the main text.

eAppendix Table 1. Association between estimated incubation period (T) and risk of death for each case of Severe Acute Respiratory Syndrome

Incubation period (T_i), per day	**Risk of Death[1] OR (95% CI)**
Cases with reported exposure dates (n=234)	
Midpoint imputation	0.89 (0.76 - 1.03)
Mean incubation time, $\overline{T}_i$	0.84 (0.68 - 1.00)
Simulated incubation times[2]	0.86 (0.71 - 1.00)
Cases in the Amoy Gardens cluster (n=308)	
Incubation period[3]	0.79 (0.67 - 0.94)

[1] The odds ratios (exp(β)) were estimated using Markov Chain Monte Carlo (10,000 runs) to estimate the coefficients (β) of the multivariable logistic regression model, with death status as the binary outcome variable and the incubation period, sex, age and occupation as predictors.

[2] 10,000 samples from the posterior distributions of the two parameters (μ, σ) and the incubation periods T for each patient were drawn using MCMC in the same algorithm in order to retain uncertainty in the incubation period in the analysis of severity.

[3] Incubation period estimated based on illness onset dates, with an assumed infection date for all cases of 21 March 2003.

eAppendix Table 2. Age stratified analysis of association between risk of death and estimated incubation period, sex and occupation.

Cases	Risk of Death[1] OR (95% CI)	
	0 - 49 years old	**≥50 years old**
Cases with reported exposure dates (n=234)	9/163[3]	16/46[3]
Incubation period[1]	0.91 (0.70 - 1.11)	0.91 (0.73 - 1.09)
Cases in the Amoy Gardens cluster (n=308)	22/236[3]	16/34[3]
Incubation period[2]	0.91 (0.65 - 0.98)	0.70 (0.65 - 0.92)

[1] The odds ratios (exp(β)) were estimated using Markov Chain Monte Carlo (10,000 runs) to estimate the coefficients (β) of the multivariable logistic regression model, with death status as the binary outcome variable and the incubation period, sex and occupation as predictors. 10,000 samples from the posterior distributions of the two parameters (μ, σ) and the incubation periods T for each patient were drawn using MCMC in the same algorithm in order to retain uncertainty in the incubation period in the analysis of severity.

[2] Incubation period estimated based on illness onset dates, with an assumed infection date for all cases of 21 March 2003.

[3] number of fatal cases/number of non-fatal cases

eAppendix Table 3. Association between estimated incubation period (T) using categorical variables (tertiles) and risk of death for each case of Severe Acute Respiratory Syndrome

Factors	Risk of death, OR (95% CrI) in cases with reported exposure dates (n=234)	Risk of death, OR (95% CrI) in cases of the Amoy Gardens cluster (n=308)[2]
Incubation period below 1st tertile (shortest)[1] (reference group)	1.00	1.00
Incubation period between 1st and 2nd tertile[1]	0.68 (0.22 - 2.11)	0.18 (0.05 - 0.63)
Incubation period above 2nd tertile (longest)[1]	0.57 (0.17 - 1.73)	0.07 (0.02 - 0.27)

[1] The odds ratios (exp(β)) were estimated using Markov Chain Monte Carlo (10,000 runs) to estimate the coefficients (β) of the multivariable logistic regression model, with death status as the binary outcome variable and the incubation period, age, *sex and occupation as predictors. The tertiles were (2.3, 5.0 days) and (2.8, 5.8 days) for the patients with exposure data and the Amoy garden cohort, respectively*

[2] Incubation period estimated based on illness onset dates, with an assumed infection date for all cases of 21 March 2003.

eAppendix Figure 1. Distribution of illness onset dates for all cases determined through epidemiologic investigations to form the Amoy Gardens cluster (n=331). The putative super spreading event happened on the 21st March 2003. The two vertical lines delimit the period of selection for the patients included in the main analysis. The dotted line indicates a lognormal distribution for the incubation period estimated on the separate subset of cases with exposure information as shown in Figure 1A in the main text (i.e. incubation period with mean 4.7 days, standard deviation 4.6 days).