Association between Severity of MERS-CoV Infection and Incubation Period

Victor Virlogeux, Minah Park, Joseph T. Wu,[1] Benjamin J. Cowling[1]

We analyzed data for 170 patients in South Korea who had laboratory-confirmed infection with Middle East respiratory syndrome coronavirus. A longer incubation period was associated with a reduction in the risk for death (adjusted odds ratio/1-day increase in incubation period 0.83, 95% credibility interval 0.68–1.03).

The incubation period of an infectious disease is the time from the moment of exposure to an infectious agent until signs and symptoms of the disease appear (*1*). This major biological parameter is part of the case definition and is used to determine duration of quarantine and inform policy decisions when mathematical modeling is used (*2*). Incubation periods vary from person to person, and their distribution tends to be right-skewed and unimodal (*3*). Variability in incubation periods for infection with Middle East respiratory syndrome coronavirus (MERS-CoV) has been described (*4–8*). Previous studies have not examined whether the length of the incubation period in a person has any correlation with subsequent clinical outcomes.

In 2015, South Korea had the largest outbreak of MERS-CoV infections outside the Arabian Peninsula (*6*). In a previous study, we reported that patients who died of severe acute respiratory syndrome (SARS) coronavirus infection had a shorter incubation period compared with infected patients who survived (*9*). The objective of this study was to examine the association between severity of MERS-CoV illness and length of incubation period.

The Study

We retrieved publicly available data from the Korea Center for Disease Control and Prevention, the Korean Ministry of Health and Welfare, the World Health Organization, and local news reports in South Korea to compile a list of all confirmed cases that had been reported by July 26, 2015 (*6*). Exposure data were available for 109 (64%) of 170 patients. For most cases, information on exposure was recorded as intervals ≥2 days during which infection was believed to have occurred, rather than exact dates of presumed infection. For the subset of patients without available exposure

Author affiliations: Ecole Normale Supérieure de Lyon, Lyon, France (V. Virlogeux); The University of Hong Kong, Hong Kong, China (V. Virlogeux, M. Park, J.T. Wu, B.J. Cowling)

data, we assumed that their incubation time was 0–21 days because 21 days was the longest incubation period reported (*9,10*). Data for patients is provided in online Technical Appendix 1 (http://wwwnc.cdc.gov/EID/article/22/3/15-1437-Techapp1.xlsx).

To estimate incubation period distribution, we fitted a gamma distribution that enabled interval censoring (*6*) by using Markov Chain Monte Carlo methods in a Bayesian framework (online Technical Appendix 2, http://wwwnc.cdc.gov/EID/article/22/3/15-1437-Techapp2.pdf) (*9*). In this analysis and analyses described below, we specified flat priors for each parameter and drew 10,000 samples from the posterior distributions after a burn-in of 5,000 iterations.

To evaluate potential factors, such as age and sex, that could be associated with length of incubation period, we fitted a multiple linear regression model to the data with the log incubation period as response variable and age and sex as explanatory variables. To determine the association between incubation period and severity of disease, we first estimated the difference in mean incubation period between patients who died and those who survived. However, this analysis could not account for potential confounders. Therefore, we specified a multivariable logistic regression model in which death was the binary response variable and predictors included age, sex, and the incubation time for each patient (*9*). We performed this analysis by using an exact likelihood approach and incubation times resampled from the 10,000 posterior samples in each iteration (online Technical Appendix). All analyses were conducted by using R version 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria). Raw data and R syntax enabling reproduction of results are available from the Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.v3456).

Of 170 patients in this study, 36 (21%) died. Mean patient age was 54.6 years, and 98 (58%) were male. Patients who died were significantly older than patients who survived (68.9 years vs. 50.8 years; p<0.001). No differences regarding age, sex, and case-fatality risk were observed between patients with or without recorded exposure data. We estimated a mean incubation period of MERS-CoV in all 170 patients of 6.9 days (95% credibility interval [CrI] 6.3–7.5 days) by using a gamma distribution. Age and sex had no associations with incubation period.

The mean incubation period was 6.4 days (95% CrI 5.2–7.9 days) for 36 patients who died compared with 7.1

days (95% CrI 6.3–7.8 days) for 134 patients who survived (Figure). The difference in means was 0.62 days (95% CrI -0.99 to 2.04 days). In the multivariable logistic regression model, we found that a longer incubation period was associated with a marginally reduced risk for death (odds ratio 0.83/1-day increase in incubation period, 95% CrI 0.68–1.03/day) after adjustment for age and sex (online Technical Appendix 2 Table 2).

To examine sensitivity of our results, we also fitted the logistic regression models by using 3 categories for the incubation period. We observed similar results and a reduced risk for death associated with longer incubation periods (online Technical Appendix Table 2). Results were also consistent in the subset of 109 patients with recorded exposure intervals (online Technical Appendix Table 2).

Conclusions

We estimated the incubation period of MERS-CoV cases during the recent MERS outbreak in South Korea and found that patients who died had a shorter incubation period than patients who survived. In a previous study, we found that the length of incubation period in patients infected with SARS coronavirus was also correlated with severity of the disease, with a shorter incubation period for patients who died (9). The pathogenesis of MERS-CoV and SARS coronavirus infection is similar (11), with a rapid progression to respiratory failure and intubation occurring ≈1 week after onset of symptoms and up to 5 days earlier in MERS patients than in SARS patients (4,12). Moreover, high rates of hemoptysis were observed in patients infected with MERS-CoV, which suggests severe lung injury (4).

MERS-CoV also has higher replication rates and shows broader cell tropism in the lower human respiratory tract than severe acute respiratory syndrome coronavirus

(13). These results suggest that a shorter incubation period could be related to a higher initial infective dose and consequently to faster or greater pathogen replication. This finding could lead to a more severe disease induced by more aggressive and damaging inflammatory responses (14). Closer monitoring of patients who have a shorter incubation period could be considered during such outbreaks.

Another potential explanation for our findings is that patients with longer incubation periods were identified and infection confirmed more quickly. This improvement in time to identification and admission to a hospital led to improved prognosis. Although longer incubation periods were correlated with shorter delays from onset to laboratory confirmation, we did not find evidence of a strong mediating effect of delay from onset to laboratory confirmation on the risk for death. However, with the small sample size, there was limited statistical power to detect a small-to-moderate effect.

Our study had some limitations. Our estimates of the incubation period were based on self-reported exposure data, which could be affected by recall bias. Moreover, 61 patients (36%) included in our main analysis had missing exposure data, and inclusion in a Bayesian framework with a wide interval of 0–21 days was necessary. Both of these limitations could have reduced the statistical power of our study to identify an association. Finally, we did not have information on underlying medical conditions or the geographic location of cases, or the treatments that were given to cases, and these variables could have been associated with clinical outcomes.

In conclusion, we found an association between shorter incubation periods among patients with MERS-CoV infection and a higher risk of death subsequently, similar to the association previously reported for severe acute respiratory syndrome coronavirus (9). This association might occur because the duration of the incubation period is an early reflection of disease pathogenesis.
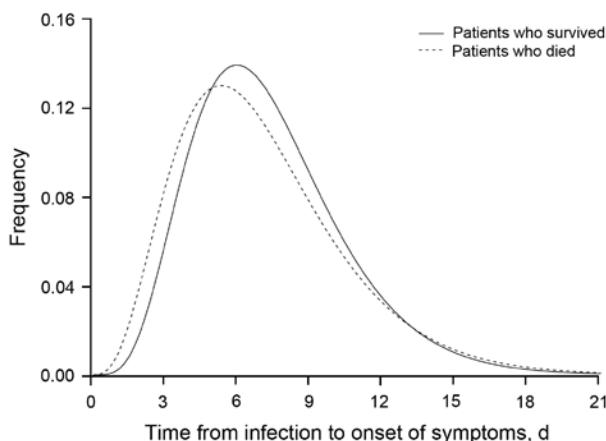
Figure. Parametric estimates of incubation period distribution for patients who died of infection with Middle East respiratory syndrome coronavirus (dashed line) and patients who survived infection (solid line), South Korea, 2015.

Dr. Virlogeux is a MD/PhD student at the University of Medicine Lyon-Est and at the Ecole Normale Superieure in Lyon, France. His research interests include infectious disease epidemiology, and the mathematical modelling of epidemic dynamics.

References

1. Fine PE. The interval between successive cases of an infectious disease. Am J Epidemiol. 2003;158:1039–47. http://dx.doi.org/10.1093/aje/kwg251
2. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DA. Incubation periods of acute respiratory viral infections: a systematic review. Lancet Infect Dis. 2009;9:291–300. http://dx.doi.org/10.1016/S1473-3099(09)70069-6
3. Sartwell PE. The distribution of incubation periods of infectious disease. Am J Hyg. 1950;51:310–8.
4. Assiri A, Al-Tawfiq JA, Al-Rabeeah AA, Al-Rabiah FA, Al-Hajjar S, Al-Barrak A, et al. Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study. Lancet Infect Dis. 2013;13:752–61. http://dx.doi.org/10.1016/S1473-3099(13)70204-4
5. Assiri A, McGeer A, Perl TM, Price CS, Al Rabeeah AA, Cummings DA, et al. Hospital outbreak of Middle East respiratory syndrome coronavirus. N Engl J Med. 2013;369:407–16. http://dx.doi.org/10.1056/NEJMoa1306742
6. Cowling BJ, Park M, Fang VJ, Wu P, Leung GM, Wu JT. Preliminary epidemiological assessment of MERS-CoV outbreak in South Korea, May to June 2015. Euro Surveill. 2015;20:7–13. http://dx.doi.org/10.2807/1560-7917.ES2015.20.25.21163
7. Cauchemez S, Fraser C, Van Kerkhove MD, Donnelly CA, Riley S, Rambaut A, et al. Middle East respiratory syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility. Lancet Infect Dis. 2014;14:50–6. http://dx.doi.org/10.1016/S1473-3099(13)70304-9
8. Centers for Disease Control and Prevention. Update: severe respiratory illness associated with Middle East respiratory syndrome coronavirus (MERS-CoV)—worldwide, 2012–2013. MMWR Morb Mortal Wkly Rep. 2013;62:480–3.
9. Virlogeux V, Fang VJ, Wu JT, Ho L-M, Peiris JS, Leung GM, et al. Brief report: incubation period duration and severity of clinical disease following severe acute respiratory syndrome coronavirus infection. Epidemiology. 2015;26:666–9. http://dx.doi.org/10.1097/EDE.0000000000000339
10. South Korean woman has MERS – a week after quarantine over. The Straits Times; 2015 [cited 2015 Oct 15]. http://www.straitstimes.com/asia/east-asia/s-korean-woman-has-mers-a-week-after-quarantine-over
11. Hui DS, Memish ZA, Zumla A. Severe acute respiratory syndrome vs. the Middle East respiratory syndrome. Curr Opin 2014;20:233–41. http://dx.doi.org/10.1097/MCP.0000000000000046
12. Peiris JS, Chu CM, Cheng VC, Chan KS, Hung IF, Poon LL, et al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. Lancet. 2003;361:1767–72. http://dx.doi.org/10.1016/S0140-6736(03)13412-5
13. Chan RW, Chan MC, Agnihothram S, Chan LL, Kuok DI, Fong JH, et al. Tropism of and innate immune responses to the novel human betacoronavirus lineage C virus in human ex vivo respiratory organ cultures. J Virol. 2013;87:6604–14. http://dx.doi.org/10.1128/JVI.00009-13
14. Ho M-S, Chen W-J, Chen H-Y, Lin S-F, Wang M-C, Di J, et al. Neutralizing antibody response and SARS severity. Emerg Infect Dis. 2005;11:1730–7. http://dx.doi.org/10.3201/eid1111.040659

Address for correspondence: Benjamin J. Cowling, World Health Organization Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Rd, Pokfulam, Hong Kong, China; email: bcowling@hku.hk

Association between Severity of MERS-CoV Infection and Incubation Period

Technical Appendix 2

Additional Details of Statistical Methods

The incubation period of an infectious disease is the time from the moment of exposure to an infectious agent until signs and symptoms of the disease appear. 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 convenient 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 in the setting of Middle East respiratory syndrome coronavirus, with each exposure interval being relatively short.

To evaluate the incubation period distribution, we compared the goodness of fit of different parametric models (gamma, lognormal, Weibull, and exponential distributions) usually used to describe the incubation period distribution of infectious diseases. We found that the gamma distribution had the best Bayesian Index Criterion value. We assumed consequently a gamma distribution with parameters (k, θ) and probability density function

$$f(t_i) = \frac{t_i^{k-1} e^{-\frac{t_i}{\theta}}}{\Gamma(k)\theta^k}$$

We assumed that the incubation period distribution had different parameters among the nonfatal cases and the fatal cases, and we consequently estimated 2 different parameters (k, θ) of the gamma distribution using Markov Chain Monte Carlo (MCMC) methods. We compared the mean incubation period between these 2 groups by using the 10,000 posterior samples of each

couple of parameters (k, θ). We then considered 2 approaches to estimate the association between the risk for death (outcome) and the incubation period (explanatory variable).

Multiple Linear Regression

We evaluated the potential association between the length of the incubation period and the age and sex of patients in both subgroups (fatal and nonfatal cases) by using a multiple linear regression approach within a Bayesian framework, and we did not find a significant association (Table 3).

Approach 1: Exact Likelihood Approach

Let f and F be the pdf and cdf of the incubation period, assumed to be gamma distributed with parameters k and θ and stratified by clinical outcome (fatal and nonfatal cases). Let P be the probability of death, which we assume to be dependent on age (g), sex (s) and incubation period (x) as in logistic regression:

(1)

$$P(g, s, x) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 g + \beta_2 s + \beta_3 x)]}$$

Depending if the case i had an exact exposure date (A_1) or an interval of exposure (A_2), we defined the probability of death q_i. If case i is in A_1, then the probability of death is simply $q_i = P(g_i, s_i, x_i)$. If case i is in A_2, then the probability of death is

$$q_i = \int_{x_i^L}^{x_i^U} P(g_i, s_i, x) \frac{f(x|k_i, \theta_i)}{F(x_i^U|k_i, \theta_i) - F(x_i^L|k_i, \theta_i)} dx$$

where $[x_i^L, x_i^U]$ is the range of incubation period for case i and where (k_i, θ_i) is the couple of parameters of the gamma distribution depending if the case i belongs to the fatal or nonfatal cases group.

We estimated $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, k_{fc}, \theta_{fc}, k_{nfc}, \theta_{nfc})$ simultaneously using MCMC and the following likelihood:

(2)

$$L(\theta) = \prod_{i \in A_1} f(x_i|k_i, \theta_i) \prod_{i \in A_2} [F(x_i^U|k_i, \theta_i) - F(x_i^L|k_i, \theta_i)] \prod_i q_i^{d_i}(1 - q_i)^{1-d_i}$$

where $d_i = 1$ if case i died from the disease and 0 otherwise; (k_{fc}, θ_{fc}) and (k_{nfc}, θ_{nfc}) are the 2 parameters of the gamma distribution for the fatal and the nonfatal cases, respectively.

Approach 2: Resampling Approach

We also defined a logistic regression model by using incubation times resampled from the 10,000 posterior samples. This approach enabled us to simulate the distribution with imputed values for individual incubation periods, which was particularly useful for an analysis in which we stratified incubation periods into tertiles. In general, the likelihood based approach might be preferred to this simulation approach, and we presented the simulation approach results as sensitivity analyses.

In this approach, the probability of death was similarly defined as in equation (1) and for each patient with interval-censored exposure data, we estimated 10,000 posterior samples for the incubation time by using MCMC, and we used the same likelihood as defined in equation (2), but where $q_i = P(g_i, s_i, x_i)$ for all cases, using the resampled incubation time for patients with interval-censored data.

Bayesian Framework

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

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

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 noninformative 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:

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

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

Technical Appendix Table 1. Characteristics of patients with cases of infection with MERS-CoV, South Korea*

Patient characteristics	Fatal cases	Nonfatal cases	Overall	p value
All patients				
Sample size, no. (%)	36 (21)	134 (79)	170 (100)	–
Mean ± SD age, y	68.9 ± 10.0	50.8 ± 15.4	54.6 ± 16.2	<0.001
Male sex, no. (%)	24 (67)	74 (55)	98 (58)	0.297
Mean incubation period, d (95% CrI)	6.4 (5.2–7.9)	7.1 (6.3–7.8)	6.9 (6.3–7.5)	–
Patients with recorded exposure intervals				
Sample size, no. (%)	26 (24)	83 (76)	109 (100)	–
Mean ± age, y	68.6 ± 10.0	50.4 ± 14.6	54.8 ± 15.7	<0.001
Male sex, no. (%)	18 (69)	47 (57)	65 (60)	0.361
Mean incubation period, d (95% CrI)	6.4 (5.2–8.0)	7.1 (6.4–7.8)	6.9 (6.3–7.5)	–

*MERS-CoV, Middle East respiratory syndrome coronavirus; –, not applicable; CrI, credibility interval.

Technical Appendix Table 2. Factors associated with risk for death from infection with MERS-CoV, South Korea*

	Risk for death,† OR (95% CI)	
Factors	All patients, n = 170	Patients with recorded exposure intervals, n = 109
Approach 1: continuous incubation period using exact likelihood		
Incubation period (continuous)	0.83 (0.68–1.03)	0.91 (0.75–1.10)
Age, y	1.11 (1.07–1.16)	1.13 (1.09–1.19)
Sex, M vs. F	2.24 (0.89–6.00)	3.15 (0.98–10.10)
Approach 2: continuous incubation period using resampling method		
Incubation period‡ (continuous)	0.81 (0.66–0.98)	0.91 (0.73–1.12)
Age, y	1.11 (1.08–1.16)	1.15 (1.09–1.23)
Sex, M vs. F	1.89 (0.73–5.32)	3.56 (1.02–13.86)
Approach 2: incubation period split into tertiles		
Incubation period‡		
Less than 1st tertile (shortest)§ (reference group)	1.00	1.00
1st–2nd tertile§	0.55 (0.20–1.48)	0.67 (0.07–3.25)
Greater than 2nd tertile (longest)§	0.26 (0.09–0.91)	0.62 (0.11–3.11)
Age, y	1.12 (1.08–1.16)	1.14 (1.08–1.20)
Sex, M vs. F	2.27 (0.84–7.15)	3.04 (0.91–10.84)

*MERS-CoV, Middle East respiratory syndrome coronavirus; OR, odds ratio.
†Coefficients exp(β) of the logistic regression were estimated by using Markov Chain Monte Carlo methods (10,000 runs) with incubation period as outcome variable and age and sex as predictors. Moreover, 10,000 samples from posterior distributions of incubation periods T for each patient estimated were used in the logistic regression model.
‡10,000 samples of the incubation periods T for each patient were drawn by using Markov Chain Monte Carlo methods.
§Tertiles were 5.1 and 8.0 days for all patients and 5.2 and 8.1 days for patients with exact exposure dates, respectively.

Technical Appendix Table 3. Factors associated with incubation period in fatal and nonfatal cases of infection with MERS-CoV, South Korea*

Factor	Coefficient β (95% CrI)†	
	Fatal cases, n = 36	Nonfatal cases, n = 134
Age	−0.06 (−0.16 to − 0.04)	0.02 (−0.01 to − 0.04)
Sex, M vs. F	−1.02 (−3.38 to − 1.51)	−0.34 (−1.25 to − 0.58)

*MERS-CoV, Middle East respiratory syndrome coronavirus; CrI, credibility interval.
†Coefficients (β) of multiple linear regression were estimated by using Markov Chain Monte Carlo methods (10,000 runs) with incubation period as outcome variable and age and sex as predictors. Moreover, 10,000 samples from posterior distributions of incubation periods T for each patient estimated were used in the multiple regression model.