University of California, Berkeley
U.C. Berkeley Division of Biostatistics Working Paper Series

Year 2003 *Paper* 133

Temporal Stability and Geographic Variation in Cumulative Case Fatality Rates and Average Doubling Times of SARS Epidemics

Alison P. Galvani* Xiudong Lei†
Nicholas P. Jewell‡

*Department of Molecular and Cell Biology, University of California, Berkeley, agalvani@nature.berkeley.edu

†Division of Biostatistics, School of Public Health, University of California, Berkeley

‡Division of Biostatistics, School of Public Health, University of California, Berkeley

http://biostats.bepress.com/ucbbiostat/paper133

Temporal Stability and Geographic Variation in Cumulative Case Fatality Rates and Average Doubling Times of SARS Epidemics

Alison P. Galvani, Xiudong Lei, and Nicholas P. Jewell

Abstract

We analyze temporal stability and geographic trends in cumulative case fatality rates and average doubling times of severe acute respiratory syndrome (SARS). In part, we account for correlations between case fatality rates and doubling times through differences in control measures. We discuss factors that may alter future estimates of case fatality rates. We also discuss reasons for heterogeneity in doubling times among countries and the implications for the control of SARS in different countries and parameterization of epidemic models.

Temporal stability and geographic variation in cumulative case fatality rates and average doubling times of SARS epidemics

Concern over the emergence of severe acute respiratory syndrome (SARS) persists as the epidemic continues despite control efforts. As of May 12, the World Health Organization (WHO) had reported 7447 cases of SARS, with 552 deaths, scattered over more than 30 countries. The most affected locations have been China, Hong Kong, Singapore, Viet Nam, Taiwan and Canada, which we focus on here. Doubling times and case fatality rates are fundamental to the epidemiology and potential public health impact of SARS. Case fatality rates (CFRs) are defined as the proportion of cases that are fatal. The doubling time is the period required for the epidemic to double, and is thus a measure of the rate of spread of disease. Doubling times also indicate the magnitude of control effort required to curtail spread. It is important to stress that doubling times change substantially over the course of an epidemic so that it is inappropriate to use current estimates to extrapolate into the future.

Case fatality rates of SARS have typically been estimated by dividing the number of fatalities by the total number of cases. This method is sufficient for an advanced epidemic. However, the method is not accurate at an early stage of an epidemic, particularly when the time from infection to recovery/death is not short relative to the duration of the epidemic, as is currently the case for SARS. The method underestimates case fatality, because it does not take into account that a proportion of individuals who are currently infected will die from the disease. A more accurate method is to divide the number of deaths by the total number of deaths + recovered. Applying this method to

publicly available WHO data (1), stability in the cumulative CFR estimates emerge, with relatively constant CFRs within a country (aside from Taiwan which is in the very earliest stage of its epidemic), but with considerable heterogeneity among countries (Figure 1A).

Overall, CFR averaged over all countries has grown from 10.4% on April 21 to 14.7% on May 12 (largely due to the sudden jump in CFR in China and Taiwan over this period). In countries with few deaths this estimate of CFR may be a slight overestimate if time from infection to death tends to be less than time to recovery. However, recent cohort data from Hong Kong (2) suggest the opposite implying that our crude estimates of CFR may still be underestimates. Such inaccuracies are however unlikely to modify the general comparison of CFRs across countries discussed below. Nevertheless, caution must be used in comparing CFRs across countries since there may be differences in the various surveillance systems that report cases and the number recovered.

Note that Figure 1A does not directly provide information on whether the CFR shows temporal trends in any country as it plots the average CFR since the beginning of the epidemics. Unfortunately, the publicly available WHO data does not permit estimation of a CFR over time since cases reported in one time period are not linked to recoveries at the same or future time. Determination of factors, including date of infection, that influence fatality rates await detailed analyses of cohort data on infected individuals.

We identify an inverse relationship between the average CFR and the average doubling time for different countries (Figure 1B). (The average doubling time is a cumulative measure reflecting average growth from the beginning of reported data, and is

estimated by the length of a time period divided by the $\log_2$ of the relative growth in reported cases over the same period—see Appendix).

This relationship is probably generated by the influence of the efficacy of control policy affecting both parameters, rather than a reflection of different characteristics of viral infectiousness and virulence across epidemics. Rapid hospitalization of infectious individuals is likely both to reduce the CFR and to increase the doubling time (by reducing the spread of SARS). Consistent with this explanation is the successful containment of a sizable epidemic in Viet Nam and the relatively low CFR and long doubling time there. In contrast, Canada has the highest CFR and shortest doubling time (not counting Taiwan where the CFR has yet to reach a steady state). Stochasticity in personal contacts plays a key role during the invasion phase of the epidemic. In Toronto it resulted in a second outbreak after public health officials had thought that SARS had been controlled. Transmission also occurred in Toronto before public awareness of SARS was widespread, resulting in a delay in the hospitalization of the first few cases that in turn facilitated transmission and may have elevated fatality rates (3). In other words, the means of disease introduction may be important in determining early doubling times and CFR values. These factors, combined with small sample sizes, may be the cause of particularly high CFRs from April 10 to 15th in Canada; it remains to be seen whether high CFRs can be attributed to alternative explanations.

Variation in CFRs among countries will arise from differences in intensity and rapidity of medical care, age-structure (older infected patients are more likely to die, (2)) and other factors such as coinfection. For example, the high prevalence of coinfection with other respiratory diseases, such as *Chlamydia pneumoniae* (4), *C. psittaci* and

paramyxoviruses in China could increase the CFR there. Likewise, should it spread in Africa, SARS could have a devastating effect, given the high prevalence of tuberculosis and HIV/AIDS, although the impact of such concomitant infections on SARS susceptibility and outcome remains to be studied.

Estimates of CFR may change as polymerase chain reaction (PCR) assays become more widely used in diagnosis (5). Diagnostic tests could identify mild cases of SARS currently escaping notification. This would increase our estimate of the size of the epidemic in terms of number of cases, but would reduce estimates of the CFR. In the other direction, PCR tests might eliminate SARS as a cause in some suspected cases. Ultimately, accurate estimates of population distributions of parameters reflecting the clinical course of disease will be best provided by follow-up of clearly defined cohorts of infected individuals identified by appropriate diagnostic procedures.

As an epidemic declines, the doubling time increases. Variation in doubling time among countries probably arises from variation in both transmission rates and control efforts (Figure 1C). Transmission rate (with units of $time^{-1}$) is determined by the expected number of susceptibles each infectious individual contacts during a unit period of time in their infectious periods, and by the probability of disease transmission per contact. High-density population centers, crowded public transport and hospital waiting rooms, enhance number of contacts, while personal hygiene affects the probability of transmission per contact. In all countries, seasonal effects may also play a substantial role with increased spread in winter as compared to late spring and summer months, as in the case of influenza transmission.

In Viet Nam, the doubling time increased over the time that the epidemic was being controlled (Figure 1C). The dramatic drop in doubling time in China in early April corresponds to a change in reporting practices (Figure 1C). Similarly, in the US, a shift in the definition of a SARS case to correspond to that recommended by the WHO complicates estimation of doubling time. However, the US appears to have a longer doubling time, because most cases are due to seeding from travel to Asia, with few cases from local chains of transmission.

Epidemic models may provide a framework for evaluating alternative control measures. Central to the accurate parameterization of epidemic models is the reproductive ratio, R_0, which is the average number of secondary cases generated by one initial infection in a susceptible population, and in the absence of control measures (6). R_0 defines a threshold that determines whether an infection is likely to spread. If R_0 is less than one, each infection will not replace itself, on average, and the disease will likely die out, although in such cases spatial dynamics, latency, and stochastic variation may all contribute to localized flare-ups of the disease that may persist for a considerable length of time. Thus, R_0 also defines the level of intervention required to contain an epidemic. The doubling rate can be used to calculate R_0 given that $R_0 = 1 + (\gamma + \alpha)\ln 2/\tau$ where γ is the duration of the incubating period, α is the duration of the symptomatic period and τ is the doubling time (6). Accurate characterization of the incubation and symptomatic periods is essential to the translation of doubling times to R_0. Typical estimates of the incubation period for SARS range from about 2 to 10 days, with both a mean and a median of about 5 days (3; 7), while the symptomatic period has a mean (± SD) of 16 ± 8 days (3; 7). Recent data from Hong Kong (2) suggests somewhat longer incubation on

average. However, severe infections may be over-represented in current estimates. Published estimates have been based largely on individuals who have received intensive medical treatment, another factor that may affect the symptomatic period. At this point in any of the epidemics, we are reluctant to use this approach for calculation of R_0 from doubling times since the latter is confounded by evolving control policies (e.g. Hong Kong, Toronto); the most natural epidemic (in Guangdong) offers the least complete data.

Figure 2 plots the reported case counts in China together with an exponential curve fitted to a smooth version of the counts (to allow for the discreteness in reports in early April). The estimated doubling time from this curve is 16.2 (which closely matches the May 3 value for China in Fig 1C which is 16.3—the May 12 doubling time is now 17.7 as the growth in counts has declined in the 9 days since May 3); note that the curve suggests that 502 cases existed in China on March 17 (with a 95% confidence interval of (468, 538)) (Appendix), consistent with under-reporting at that time. We again stress that control measures, evolving contact patterns, stochastic effects and potential acquired immunity will all impact this doubling time (equivalently, the growth in case counts) in the future, and will ultimately lead to a flattening of the growth observed to date. The lack of fit of a simple exponential curve can be already seen in the last week of data from China.

This rapid growth in China suggests an urgent need to control the epidemic in Asia before it gains further momentum. Containment of an outbreak at an early stage affords a greater chance of success than does a later response, and the former clearly puts less strain on the health care system. Isolation of cases, infection-control measures in hospitals, and vigilant surveillance at community and population levels are imperative.

Failing this, SARS could become endemic in China, particularly if it evolves antigenically to evade pre-existing immunity, such that recovered patients could be reinfected, as is the case for influenza (8). In this eventuality, international travel would continually seed new cases in other parts of the world. SARS reaffirms what we have previously learned from other infectious diseases, namely that epidemic control is a global concern and not the problem of one or a few nations.

Acknowledgements

We thank A. Reingold for discussions.

Biography of authors

Dr. Alison Galvani is a Miller Research Fellow at University of California, Berkeley. She received her PhD from the University of Oxford. Her primary research interests are focused on the evolution and epidemiology of infectious diseases, including influenza. Dr Nicholas P. Jewell is Professor of Biostatistics and Statistics in the Division of Biostatistics at School of Public Health at University of California, Berkeley. He developed and applied statistical techniques to early phases of the HIV epidemic. He was Vice Provost of the University of California, Berkeley from 1994 to 2000. He has returned to full-time research in the development of biostatistical techniques for epidemiological data analysis. Dr Jewell is an editor of *Biometrika*. Xiudong Lei is a doctoral student in Biostatistics at the University of California, Berkeley.

References

1. Cumulative number of reported cases of severe acute respiratory syndrome (SARS). Geneva: World Health Organization, 2003 (Accessed May 12, 2003 at http:// www.who.int/csr/sarscountry/en/)

2. Donnelly, CA, Ghani, AC, Leung, GM, Hedley, AJ, Fraser, C, Riley, S, Abu-Raddad, LJ et al. Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. Lancet online May 7, 2003; http://image.thelancet.com/extras/03art4453web.pdf.

3. Poutanen, SM, Low, DE, Henry, B, Finkelstein, S, Rose, D, Green, K, Tellier, R et al. Identification of severe acute respiratory syndrome in Canada. N. Eng. J. Med. 2003; 348.

4. Koh, WP, Taylor, MB, Hughes, K, Chew, SK, Fong, CW, Phoon, MC, Kang, KL, at al. Seroprevalence of IgG antibodies against *Chlamydia pneumoniae* in Chinese, Malays and Asian Indians in Singapore. Int. J. Epid. 2002; 31:1001-7.

5. Dorsten, C, Gunther, S, Preiser, W, Werf, S, Brodt, HR, Becker, S, Rabenau, H et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Eng. J. Med. 2003; 348.

6. Anderson, RM, May, RM. *Infectious Diseases of Humans* (Oxford University Press, Oxford, 1991)

7. Tsang, KW, Ho, PL, Ooi, GC, Yee, WK, Wang, T, Chan-Yueng, M, Lam, WK et al. A cluster of severe acute respiratory syndrome in Hong Kong. N. Eng. J. Med. 2003; 348.

8. Ferguson, NM, Galvani, AP, Bush, RM Ecological and immunological determinants of influenza evolution. Nature 2003;422:428-33.

9. Cleveland, WS. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 1979;74:829-36.

Appendix

Definition of doubling time

The average doubling time over the interval from time t_0 to time t_1 is simply

$$\frac{t_1 - t_0}{\log_2\left(N_1 / N_0\right)}$$

where N_1 and N_0 are the number of cases at times t_1 and t_0, respectively. The units correspond to those used to measure the interval length t_1-t_0. As an example, if $N_1 = 2N_0$, then the average doubling time is exactly t_1-t_0. In Figures 1B and 1C, the time t_0 (in days) is always taken to be the earliest time where case counts are available in the WHO data.

Analysis of SARS data of China

Figure 2 was obtained by (i) first applying a lowess smooth (9) to the observed case counts, (ii) applying least squares linear fit of the logarithm of the smoothed case counts against time, and (iii) transforming the fitted line back to the original scale . In particular, in (ii) the estimated line is

log(case) = 6.176 + 0.0427*time

(where time in days ranges inclusively from the value 1 (March 17) to 57 (May 12). The squared correlation coefficients are: R^2 =0.98 (for the original cumulative case counts and the lowess smoothed counts), R^2 =0.95 (for the lowess smoothed counts and the

fitted counts), and R^2 =0.91 (for the original cumulative case counts and the fitted counts).

R square for the lowess smoothed case counts and the fitted counts (from the exponential curve) is 95.39%

Prediction of cases with 95% confidence interval for March 17, 2003

The predicted confidence interval for Y(new) is as following:

$$\frac{Y_{(new)} - \hat{Y}}{\sqrt{Var(pred)}} \rightarrow t_{(n-2)}, \quad Var(pred) = MSE * \left[1 + \frac{1}{n} + \frac{(X_{new} - \overline{X})^2}{\sum (X_i - \overline{X})^2}\right]$$

The estimates and confidence intervals are then transformed back to the original scale.

On March 17, 2003, the estimated case is 502, with a 95% confidence interval (468, 538)—note that this confidence interval is for the actual number of cases on March 17.

Figure 1A: Time series of cumulative case fatality rate.

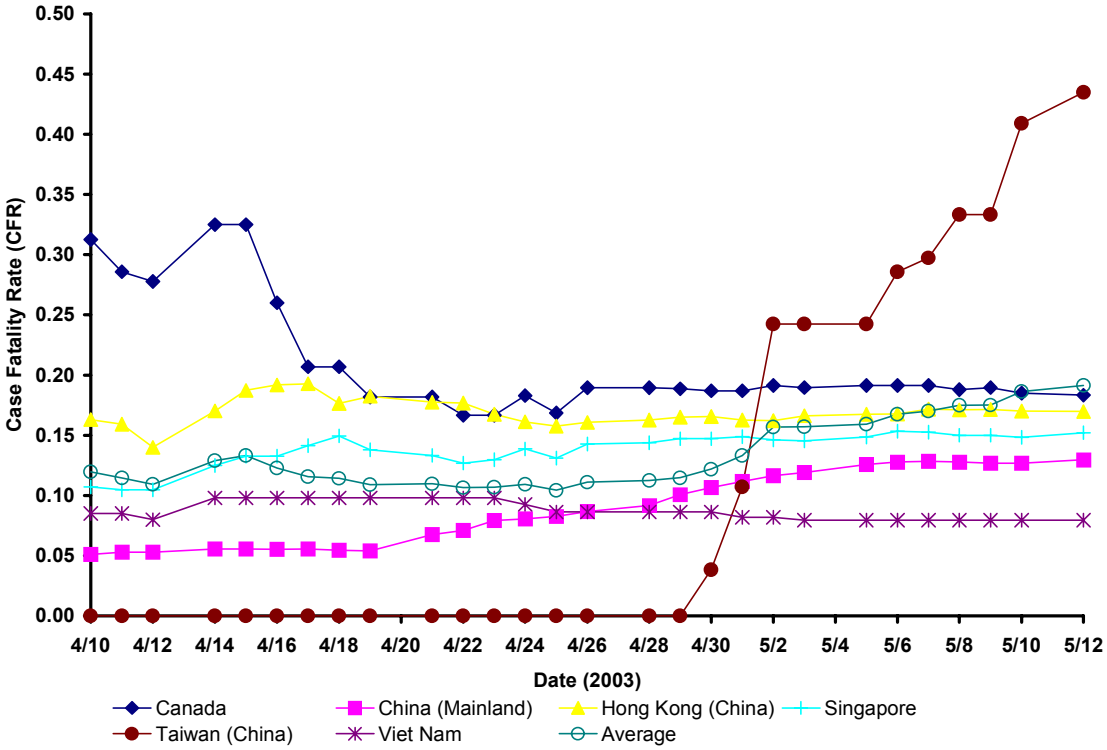

Figure 1B: Cumulative case fatality rate against average doubling time as of May 12, 2003

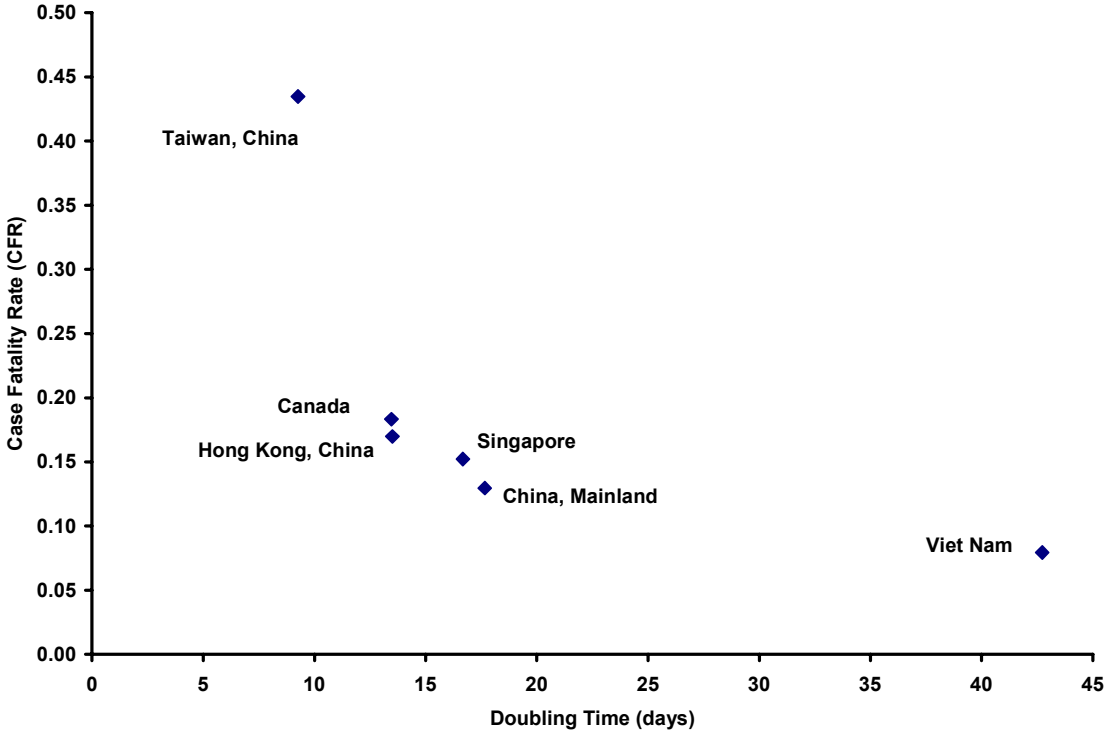

Figure 1C. Time series of log of average doubling time

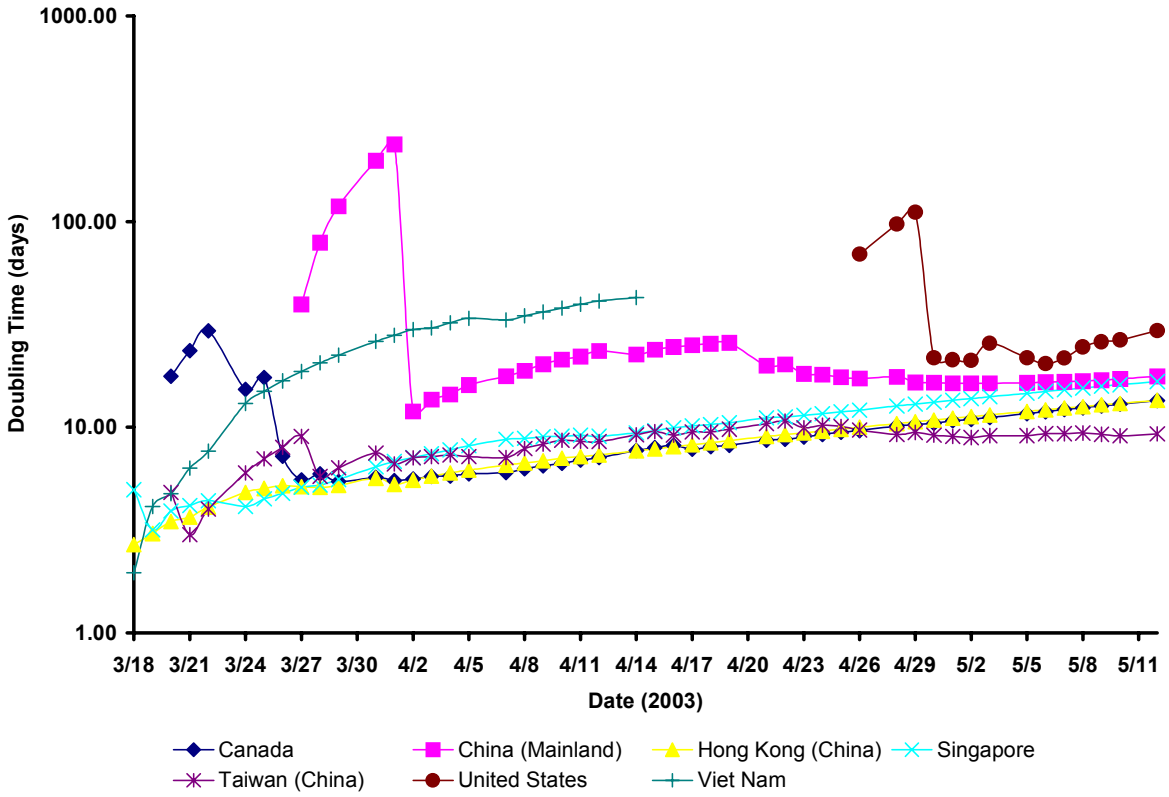

Figure 2: Observed and expected cumulative number of SARS cases in China