RESEARCH METHODS & REPORTING

Assessing the severity of the novel influenza A/H1N1 pandemic

Tini Garske, Judith Legrand, Christl A Donnelly, Helen Ward, Simon Cauchemez, Christophe Fraser, Neil M Ferguson, Azra C Ghani

FAST TRACK

A major concern about the emergence of the novel strain of influenza A/H1N1 is the severity of illness it causes. **Tini Garske and colleagues** propose methods to obtain accurate estimates of the case fatality ratio as the pandemic unfolds

EDITORIAL by Anderson

MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London W2 1PG

Correspondence to: T Garske
t.garske@imperial.ac.uk

Accepted: 8 July 2009

Cite this as: *BMJ* 2009;339:b2840
doi: 10.1136/bmj.b2840

The World Health Organization's declaration of a pandemic of the novel influenza A/H1N1 virus raises questions about the potential morbidity and mortality. By 10 July 2009, nearly 100 000 cases had been reported worldwide; however, most deaths (429 in total) have been reported in the American continents (the US, Mexico, Argentina, and Canada), with smaller numbers in other countries including the United Kingdom.[1] At first sight, the data seem to imply that this new virus is relatively mild, with case fatality ratios around 0.5%, similar to the upper range of that seen for seasonal influenza[2] and relatively low hospitalisation ratios. However, the case fatality ratio seems to vary substantially between countries, and deaths have occurred in much younger people than is the case for seasonal influenza.[3 4]

There are many reasons why simple interpretations of these crude figures at the beginning of a pandemic may be misleading both in terms of assessing severity and in making comparisons between countries. Here, we discuss some of the important mechanisms resulting in biases, propose study designs and associated statistical methods to estimate the case fatality ratio given these limitations, and show their strengths using simulated data. The two main sources of bias in estimates of the case fatality ratio we consider stem from shifts in case ascertainment (over time, efforts may become more focused on the most severe cases, leading to an overestimation of the case fatality ratio) and from the inevitable delay between symptom onset and death, which in the early phase of the epidemic can lead to underestimation of the case fatality ratio if it is not adjusted for.

Case ascertainment—what are the numerator and denominator?

A natural definition for the case fatality ratio is the ratio of the total number of deaths from a disease divided by the total number of cases. In a fully ascertained (and complete) epidemic, this simple method works perfectly. However, in most infectious diseases there is underascertainment of cases as people who are asymptomatic or have mild infection will be less likely to present to health care, and if they do present they will be less likely to be tested and confirmed. It is therefore likely that there will be a bias towards diagnosis of more severe cases (fig 1), with the result that the case fatality ratio and other measures of severity are overestimated. Furthermore, this underascertainment will change as an epidemic matures. Initially increased awareness by patients and doctors may lead to high ascertainment, but as cases increase and systems are overwhelmed, only a proportion will be tested (potentially those with links to other confirmed cases), making it difficult to understand the scale of under-reporting.

This could introduce regional biases if, for example, testing were focused on a subset of cases in heavily affected areas but applied to all suspected cases seeking medical attention elsewhere. Taken together, these factors mean that the denominator (the total number of cases) is highly uncertain.

Furthermore, it has been shown that seasonal influenza infections can temporarily increase the risk of vascular events,[5] which might lead to excess

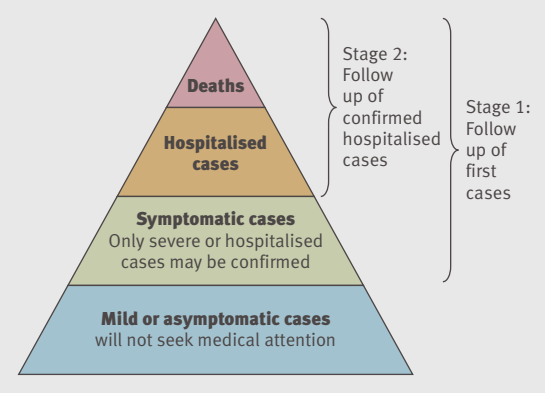

Fig 1 | Spectrum of influenza cases. Infection in patients who die or are admitted to hospital will be ascertained throughout an influenza pandemic. Case fatality ratio estimates depend on how these cases change as a proportion of total case numbers as the pandemic progresses

mortality that is not attributed to influenza, therefore underestimating the number of influenza deaths. The same effect is probably present in pandemic influenza, and there might be many more reasons why deaths that are caused by flu might not be recognised as such, particularly as many places do not have good systems of hospital surveillance. In order to get a clear picture of the severity of the novel influenza A/H1N1 it is therefore important to establish good surveillance.

The table shows the numbers of confirmed cases and deaths as well as numbers of hospital admissions for subsets of cases along with the case fatality ratios and hospitalisation ratios calculated crudely by dividing the number of deaths or hospitalisations by the number of officially reported cases up to 10 July. We also show estimates of the case fatality ratio for several countries, adjusted as proposed below for timing issues stemming from the real time nature of the estimation. Note, however, that these estimates have not been adjusted to account for the uncertainties in the ascertainment of the numerator or denominator. Details of the delay distributions used to estimate the adjusted case fatality ratio are available on bmj.com.

With the exception of Mexico, the estimated case fatality ratios are well below 1%, with those in Europe considerably lower than those in Canada and the US. Although the high case fatality ratio in Mexico could be interpreted as being due to a more virulent version of the virus or higher frequency of comorbidities resulting in more severe illness, it is more likely that case reporting there is heavily focused on the most severe cases and that the true number of cases is much higher.[12] To a lesser extent, the same phenomenon might also now be occurring in the US because the large number of cases means that testing is now biased towards severe and hospital treated cases. This is supported by the differences in hospitalisation ratios, with higher ratios in the US and Canada than in the UK, where case ascertainment has so far probably been more complete.

If we take mild unreported cases into account, the true case fatality ratios could therefore be considerably lower and comparable to that for seasonal influenza (although recent studies in ferrets suggest that morbidity might be higher[13]). However, because a large proportion of the population is probably susceptible to infection (whereas most people have some cross protective immunity to circulating seasonal influenzas), the absolute numbers of cases (and therefore also deaths) from the novel strain can be expected to be much greater than for seasonal influenza.

Although the biases due to the uncertainty in actual case numbers make comparison between countries difficult, data collected according to carefully designed protocols can limit their impact. In the UK and other countries, the first few hundred confirmed cases were closely monitored. The resulting data, as well as providing detailed clinical and epidemiological information of importance to healthcare planning, can be used to estimate the case fatality ratio, even though the number of deaths among these cases is likely to be small. The following two step procedure outlines how this can be achieved.

Firstly, in the early stages of the outbreak when all identified cases are closely followed, we can use these detailed data (with sample size n_1) to estimate the hospitalisation ratio r_H (the proportion of cases admitted to hospital). In the second stage, as the epidemic grows and full case ascertainment in the community becomes challenging, a sample of confirmed cases among hospitalised patients (sample size n_H) can be used to estimate the case fatality ratio among hospitalised cases, $r_{D|H}$. Based on these two estimates, the overall case fatality ratio can be estimated as $r_H \times r_{D|H}$. The precision of this estimate will of course depend on the precision of both estimated ratios (see bmj.com for details).

To ensure a reasonable degree of precision it is important that the number of closely monitored first cases (n_1) is of sufficient size (fig 2). For instance, in order to obtain a 95% confidence interval that covers roughly the range from 0.5 to 1.5 of the case fatality ratio, we would need around 1100 cases in the initial sample (stage 1) to estimate the case fatality ratio as soon as around 200 cases have been admitted to hospital, or 425 to estimate the ratio after 500 hospitalisations, assuming a true case fatality ratio of 0.5%

Summary of data on cases (confirmed/probable), hospitalisations, deaths, crude and adjusted case fatality ratios (%), and crude hospitalisation ratios by country or region

	No of confirmed cases*	No of confirmed deaths*	Baseline for hospital admissions†	No of hospitalised cases	Crude case fatality ratio (95% CI)	Adjusted case fatality ratio‡ (95% CI)	Crude hospitalisation ratio (95% CI)
US[6 7]	37 246	211	4566	407	0.57 (0.49 to 0.65)	0.68 (0.59 to 0.78)	8.9 (8.1 to 9.8)
Mexico[8a]	11 699	121	NA	NA	1.03 (0.86 to 1.23)	1.23 (1.03 to 1.47)	
Canada[9]	9 717	39	9717	894	0.40 (0.29 to 0.55)	0.43 (0.30 to 0.58)	9.2 (8.6 to 9.8)
UK[10a]	9 718	14	9718	335	0.14 (0.08 to 0.24)	0.24 (0.13 to 0.41)	3.4 (3.1 to 3.8)
EU[11]	13 667	16	NA	NA	0.12 (0.07 to 0.19)	0.20 (0.11 to 0.32)	

*Cases and deaths reported officially on national and international websites up to 10 July. For the US these also include probable cases.
†Sample size for which data on hospital admissions were available.
‡Adjusted for the time delays but not for uncertainties in the denominator (see equation 1 and bmj.com).

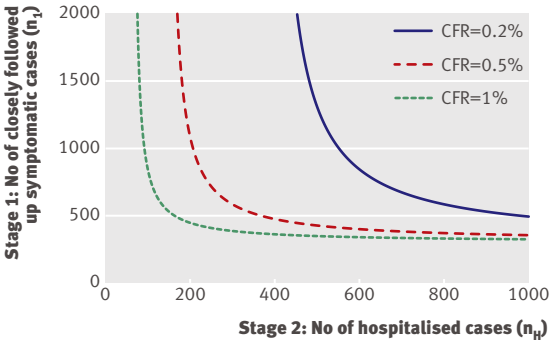

Fig 2 | Sample size of the initial closely followed cases (stage 1) and of hospitalised cases (stage 2, when full ascertainment is challenging) required to estimate the case fatality ratio (CFR) with sufficient precision to obtain a coefficient of variation of 25%, for different values of the case fatality ratio and an assumed hospitalisation ratio of 5%

and a hospitalisation ratio of 5%. For a case fatality ratio of 0.2%, the initial sample size would have to be at least 1300 cases to estimate the ratio after 500 hospitalisations, and the required precision cannot be obtained after only 200 hospitalisations, no matter how large the initial sample.

To put this into context, if we were to estimate the case fatality ratio simply by dividing the number of deaths by the number of cases, we would need detailed data on 3200 cases for a case fatality ratio of 0.5% or on 8000 cases for a ratio of 0.2%. The table shows that the number of hospitalisations in Canada and the UK are in the hundreds, whereas in the US they must be several thousand, although data on US hospitalisations are available for only a subset of cases. Note that the precision assumed here is purely illustrative. The required sample size for any other precision, or indeed at any other values of the case fatality ratio, hospitalisation ratio, and number of hospitalised cases to date can be obtained from equation 3 on bmj.com.

An important assumption underlying this approach is that the two ratios r_H and $r_{D|H}$ remain constant over time. This may not happen as the guidelines for hospitalisation of confirmed cases are likely to change over time, particularly in the move from containment to mitigation. Hospitalisation ratios may also vary between settings, depending on factors such as the availability of timely antiviral treatment and hospital beds. Hence, it is important to obtain data on the reasons for hospitalisation in order to use the hospitalisation ratio as a measure of disease severity. Alternatively, the ratio of admission to intensive care units for all identified cases r_I and the case fatality ratio among intensive care cases $r_{D|I}$ may be less subject to temporal changes in hospitalisation policies, at least as long as intensive care unit capacities are not overwhelmed. Furthermore, during the course of the pandemic, continuing studies on a representative subset of all identified cases and their outcomes, as well as cases admitted to hospital and intensive care, would ensure that monitoring of the severity of the disease remains accurate.

Even in the early stages of the epidemic full case ascertainment is an ambitious goal. If many cases are found through contact tracing rather than standard symptom based surveillance systems, it may be possible to assess the degree of mild disease. However, to be able to detect potential changes in virulence, hospitalisation ratios should be updated regularly throughout the pandemic through general syndromic surveillance and large scale testing in well defined study populations to measure the proportion of infection among people with relevant clinical symptoms. These studies need to be established prospectively and should be coupled with prospective household studies to estimate the attack rate for mild disease and serological testing to assess the extent of asymptomatic infection, such that changing patterns of virulence are detected rapidly.

Our method can readily be applied to estimate case fatality ratios stratified by characteristics such as age or comorbidity, with the expressions for case fatality ratios and variances applying separately to each stratum. This obviously decreases the numbers of cases and fatalities in each stratum and as the sample sizes needed to obtain a particular precision apply to each stratum separately, larger overall sample sizes are needed.

Delays between onset, death, and reporting

A second important source of bias arises from the delay between disease onset and knowledge of the final outcome in severe cases. Thus among the reported cases at any particular time there might be people who will eventually die but are still alive at the point of analysis (fig 3). This effect, known statistically as censoring, means that a case fatality ratio estimated crudely by dividing the cumulative number of reported deaths by the cumulative number of reported cases will be too low and will grow as the epidemic unfolds. This was observed during the severe acute respiratory syndrome (SARS) epidemic, causing concern that the virus was mutating to become more virulent.[14] Given the expectation that antigenic drift or viral reassortment with co-circulating seasonal influenzas may well change the severity of the new influenza virus over the coming months, it is especially important that these biases are minimised.

Censoring bias is particularly strong in the early phase of the epidemic, when the incidence is growing exponentially. This is because the number of cases arising this week will be greater than the number last week, in whom we are only now beginning to see deaths, increasing the degree of underestimation of the case fatality ratio. Indeed, it is only after the epidemic has peaked that this effect begins to wane.

Several methods have been proposed to account for censoring based on analysis of the SARS epidemic.[15-18] One of the simplest improvements on the crude estimate consisted of dividing the number of deaths by the total number of cases in whom the outcome was known. This seemed to work well in

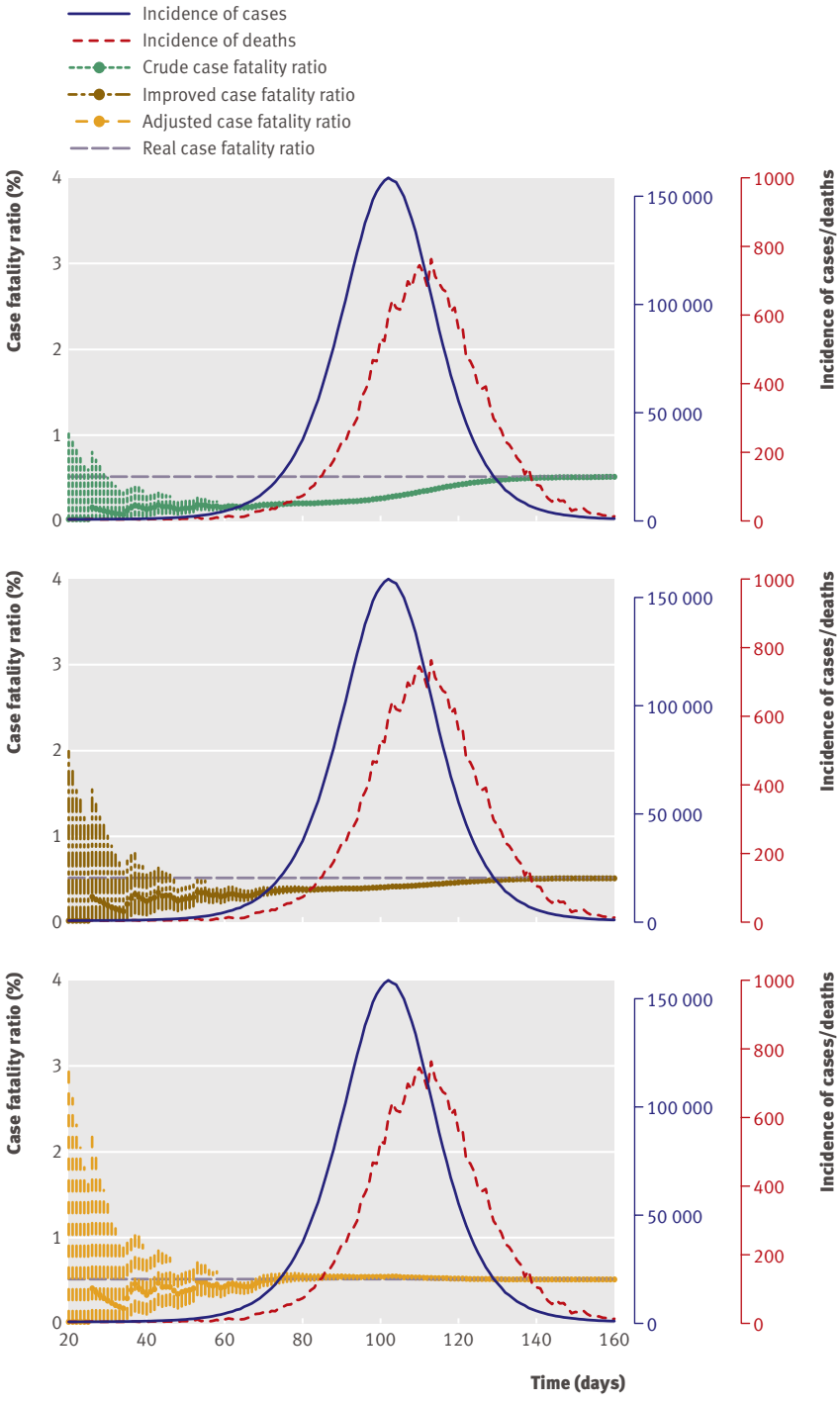

Fig 3 | Estimates of case fatality ratio with 95% exact binomial confidence intervals for a simulated epidemic using a stochastic compartmental susceptible-exposed-infectious-recovered (SEIR) type model. Top: Crude estimate obtained by dividing the reported number of deaths by the reported number of cases. Middle: Improved estimate obtained by dividing the number of deaths by the number of deaths+recovered. Bottom: Adjusted estimate obtained by dividing the number of deaths by the number of cases adjusted for the delay from symptom onset to death

analyses of the SARS epidemic. However, the times from onset to death and from onset to recovery were similar with SARS,[15] whereas for flu we would expect rapid recovery in mild cases but a longer duration of hospitalisation for people who eventually die either from viral pneumonia or secondary bacterial pneumonia. Indeed, during the 1918 pandemic the average duration from onset of symptoms to death was nine days,[19] whereas recovery from milder illness is typically 4-7 days for seasonal flu.[20-23] It is this difference that leads to a bias in the improved estimator. Furthermore, in the context of the novel influenza A/H1N1, recoveries are not readily reported and thus it is difficult to assess the degree of bias in simple estimates.

One simple method to account for these delays is to use the equation in fig 4. The cumulative distribution from time of onset to death could be taken from existing data or from past pandemics. Furthermore, it is straightforward to modify the equation to account for additional reporting delays for cases and deaths (see bmj.com) given knowledge of the reporting delay distributions.

Figure 3 shows these censoring techniques in a simulated pandemic typical of what could be expected in the UK. In the initial stages of the epidemic, when case numbers are small and subject to substantial stochastic effects, all estimates of the case fatality ratio are imprecise with wide confidence intervals and extremely variable best estimates. This variability means that bounds may be more appropriate to report than point estimates. In the exponentially growing phase, the estimates stabilise and the confidence intervals shrink as cases accrue. However, as noted, both of the crude estimators give very biased results in this stage of the epidemic. These biases only lessen once the epidemic curve is in its downturn, and finally, all the methods of estimation considered recover the true case fatality ratio with relatively narrow confidence intervals at the end of the epidemic.

The delay between symptom onset and death (note the shift between the peaks in incidence of cases versus deaths in figure 3), means that dividing the total number of deaths by the total number of cases greatly underestimates the true case fatality ratio throughout much of the epidemic but produces fairly narrow confidence intervals, giving a misleading impression. For typical influenza patterns, the second technique, taking into account only those cases whose outcome is known, also underestimates the true case fatality ratio, albeit less than the crude estimate. Our proposed estimator (see fig 4), which takes into account the duration of the delay from

$$\mathrm{CFR}(t) = \frac{D(t)}{\sum_{u=0}^{t} c(u)F(t-u)}$$

CFR(t) is the estimate of the case fatality ratio calculated on day t,
D(t) is the cumulative number of deaths reported on day t,
c(u) is the number of cases with symptom onset on day u, and
F(x) is the cumulative distribution of the time from symptom onset to death for those who will eventually die, such that F(t-u) is the probability that a case with onset on day u will already have died by the end of day t

Fig 4 | Equation to adjust for delays between onset and death

symptom onset to death, produces more reliable estimates throughout the epidemic and should be unbiased provided there is no other source of bias (such as the denominator issue) and the assumed distribution of the delay from symptom onset to death matches the true distribution. The uncertainty of the distributions can be taken into account by calculating confidence intervals of the case fatality ratio using bootstrapping methods.[24] However, if there is a bias in the true distribution of the delay from symptom onset to death because of censoring of this distribution, this method may also be biased, and hence it is important that this information is collected reliably early in the pandemic.

Conclusions

Quantitative estimates of the severity of the new influenza A/H1N1 virus are central to healthcare planning over the coming months. In addition, decisions on whether to implement social distancing measures such as school closures, will depend on a balance between the number of cases (and hence deaths) prevented and the social and economic impact of such measures. Better estimates of the true extent of infection and reliable population level estimates of the case fatality ratio will facilitate identification of risk factors between and within populations. Such information, coupled with serological surveys to determine susceptibility to the new virus, is essential for determining priority groups for vaccination when a vaccine becomes available. More data will also be needed to reliably ascertain age specific infection rates and case fatality ratios. So far, most infections have been confirmed in children and young adults, with most deaths in young adults, but it is uncertain whether this trend will persist as the virus spreads more extensively; any changes could have a considerable impact on the aggregate population level case fatality ratio.

Well designed data collection protocols as outlined here and application of relatively simple estimation methods to these data, will greatly improve our ability to obtain informative estimates of the case fatality ratio, despite the biases likely to arise in the collection of data. Critically, they will ensure that any changes in the virulence are rapidly detected so that mitigation policies are applied appropriately.

We thank Lyn Finelli, Carrie Reed, Anna Bramley, Jennifer Michalove, Ashley Fowlkes, and Mackezine Nowell for help in collecting and analysing the data. We thank US state and local health departments for the collection of linelist surveillance data. We also thank the reviewers Andrew Hayward, Sheila Bird, and Julie Morris for constructive comments.

Contributors: JL and TG developed the methods and undertook the analyses and drafted the manuscript. ACG, CAD, SC, CF, NMF, and HW contributed to the methods and drafting of the manuscript and were involved in the overall planning of the project. TG is guarantor.

This work was funded by a Medical Research Council UK Centre grant. CF was supported by the Royal Society and SC by Research Councils UK.

Competing interests: None declared.

1 World Health Organization. Situation updates—influenza A(H1N1). www.who.int/csr/disease/swineflu/updates/en/index.html.
2 US Centers for Disease Control and Prevention. *Key facts about seasonal influenza (Flu)*. www.cdc.gov/flu/keyfacts.htm, 2009.
3 Perez-Padilla R, de la Rosa-Zamboni D, Ponce de Leon S, Hernandez M, Quinones-Falconi F, Bautista E, et al. Pneumonia and respiratory failure from swine-origin influenza A (H1N1) in Mexico. *N Engl J Med* 2009 [Epub ahead of print].
4 Novel Swine-Origin Influenza A Virus Investigation Team. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med* 2009;360:2605-15.
5 Smeeth L, Thomas SL, Hall AJ, Hubbard R, Farrington P, Vallance P. Risk of myocardial infarction and stroke after acute infection or vaccination. *N Engl J Med* 2004;351:2611-8.
6 US Centers for Disease Control and Prevention. *H1N1 Flu*.www.cdc.gov/h1n1flu.
7 US Centers for Disease Control and Prevention. National linelist of influenza A(H1N1) cases, 25 June 2009.
8 Pan American Health Organisation. *Update: influenza A (H1N1) regional report (July 10 2009)*. http://new.paho.org/hq/index.php?option=com_content&task=view&id=1574&Itemid=1167.
9 Public Health Agency of Canada. Surveillance. *Pandemic (H1N1) 2009 outbreak epidemiological update*. www.phac-aspc.gc.ca/alert-alerte/swine-porcine/surveillance-eng.php.
10 Health Protection Agency. HPA weekly national influenza report, 09 july 2009 (week 28). www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1247125493338.
11 European Centre for Disease Prevention and Control. *ECDC situation reports influenza A(H1N1) infection*. www.ecdc.europa.eu/en/Health_topics/novel_influenza_virus/2009_Outbreak/
12 Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 2009:117:60-2.
13 Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, Zeng H, et al. Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. *Science* 2009:117:723-8.
14 Death rate from virus more than doubles, varying sharply by country. *New York Times* 2003 Apr 22.www.nytimes.com/2003/04/22/world/death-rate-from-virus-more-than-doubles-varying-sharply-by-country.html.
15 Ghani AC, Donnelly CA, Cox DR, Griffin JT, Fraser C, Lam TH, et al. Methods for estimating the case fatality ratio for a novel, emerging infectious disease. *Am J Epidemiol* 2005;162:479-86.
16 Jewell NP, Lei XD, Ghani AC, Donnelly CA, Leung GM, Ho LM, et al. Non-parametric estimation of the case fatality ratio with competing risks data: an application to severe acute respiratory syndrome (SARS). *Stat Med* 2007;26:1982-98.
17 Yip PSF, Lam KF, Lau EHY, Chau PH, Tsang KW. A comparison study of realtime fatality rates: severe acute respiratory syndrome in Hong Kong, Singapore, Taiwan, Toronto and Beijing, China. *J R Stat Soc Ser A Stat Soc* 2005;168:233-43.
18 Yip PSF, Lau EHY, Lam KF, Huggins RM. A chain multinomial model for estimating the real-time fatality rate of a disease, with an application to severe acute respiratory syndrome. *Am J Epidemiol* 2005;161:700-6.
19 Brundage JF, Shanks GD. Deaths from bacterial pneumonia during 1918-19 influenza pandemic. *Emerg Infect Dis* 2008;14:1193-9.
20 Hayden FG, Osterhaus A, Treanor JJ, Fleming DM, Aoki FY, Nicholson KG, et al. Efficacy and safety of the neuraminidase inhibitor zanamivir in the treatment of influenza virus infections. *N Engl J Med* 1997;337:874-80.
21 MIST Study Group. Randomised trial of efficacy and safety of inhaled zanamivir in treatment of influenza A and B virus infections. *Lancet* 1998;352:1877-81.
22 Nicholson KG, Aoki FY, Osterhaus A, Trottier S, Carewicz O, Mercier CH, et al. Efficacy and safety of oseltamivir in treatment of acute influenza: a randomised controlled trial. *Lancet* 2000;355:1845-50.
23 Treanor JJ, Hayden FG, Vrooman PS, Barbarash R, Bettis R, Riff D, et al. Efficacy and safety of the oral neuraminidase inhibitor oseltamivir in treating acute influenza: a randomized controlled trial. *JAMA* 2000;283:1016-24.
24 Gigli A, Verdecchia A. Uncertainty of AIDS incubation time and its effects on back-calculation estimates. Stat Med 2000;19:175-89.
25 Armitage P, Berry G, Matthews JNS. *Statistical methods in medical research*. 4th ed. Oxford: Blackwell, 2002.
26 Secretaria de Salud Mexico. *Situación actual de la epidemia*. http://portal.salud.gob.mx/descargas/pdf/influenza/situacion_actual_epidemia_250609.pdf.