Published by Oxford University Press on behalf of the International Epidemiological Association
© The Author 2007; all rights reserved. Advance Access publication 17 May 2007

International Journal of Epidemiology 2007;**36**:881–889
doi:10.1093/ije/dym071

Estimates of the reproduction numbers of Spanish influenza using morbidity data

Emilia Vynnycky,[1]* Amy Trindall[1,3] and Punam Mangtani[2]

Accepted 20 March 2007

Background There have been several studies of the transmissibility of the 1918 (Spanish) influenza virus, which has attributed to >20 million deaths. Many of the analyses to date have involved fitting predictions from a transmission model to the observed epidemic curves from different settings.

Methods Using morbidity data from cities in Europe and America and from confined settings during the 1918 influenza pandemic, we contrast the use of several different methods based on the growth rate and final size of the epidemic, which do not rely on transmission models, to estimate the effective and basic reproduction numbers.

Results The effective reproduction number (the average number of secondary infectious cases produced by a typical infectious case in a given population) for the 1918 influenza virus was in the range 1.2–3.0 and 2.1–7.5 for community-based and confined settings, respectively.

Conclusions Assuming further that 30 and 50% of individuals were immune to Spanish influenza after the wave in April 1918 and the first subsequent wave, respectively, these findings imply that, in a totally susceptible population, an infectious case could have led to 2.4–4.3 and 2.6–10.6 cases in community-based and confined settings, respectively. These findings for community-based populations confirm the relatively low transmissibility of the 1918 (Spanish) influenza virus, which has been found by other studies using alternative data sources and methods.

Keywords Modelling, reproduction numbers, Spanish influenza

Introduction

The recent emergence of the highly pathogenic H5N1 virus among poultry and ducks and the clinical, epidemiological and laboratory indications among the human cases attributed to H5N1 to date, has led to fears of an imminent influenza pandemic similar in impact to that of the Spanish influenza pandemic,[1] which is reported to have caused more deaths than the first World War.[2]

The impact of any pandemic and its potential for control depends greatly on the transmissibility of the causal pathogen and its ability to lead to infectious cases. This is usually described using its basic reproduction number (R_0), defined as the average number of secondary infectious cases resulting from a typical infectious case in a totally susceptible population.

There are several methods for estimating the R_0 and most of the studies to date for Spanish influenza have used similar methods, involving the development of a transmission model and fitting predictions of the epidemic curve to the available data.[3–5] The conclusions from these studies have been very similar, i.e. the R_0 was approximately 3–4.

We here illustrate the application of several different methods, which do not rely on the development of transmission models, to estimate the reproduction numbers for Spanish influenza for a variety of different settings, using data from the first few weeks of the epidemic and assumptions about the serial interval.

[1] Modelling and Economics Unit, Health Protection Agency Centre for Infections, 61 Colindale Avenue, Colindale, London, NW9 5HT, UK.

[2] Infectious Disease Epidemiology Unit, Department of Infectious and Tropical Diseases, Keppel Street, London, WC1E 7HT, UK.

[3] Public Health Directorate, West Essex Primary Care Trust, 16th Floor, Terminus House, Terminus Street, West Essex, CM20 1EX, UK.

* Corresponding author. Modelling and Economics Unit, Health Protection Agency Centre for Infections, 61 Colindale Avenue, Colindale London, NW9 5HT. E-mail: emilia.vynnycky@hpa.org.uk.

Materials and methods

Data sources

Data were obtained after a review of the English language literature to identify suitable studies presenting daily or weekly data on the incidence of influenza during 1918–19. Four types of study settings were identified from the spring of 1918, when an increase in influenza incidence was first observed, namely: confined populations [the San Quentin prison (USA) and Australian boats], communities in Maryland, USA, where cases were ascertained actively through house-to-house surveys and several cities in Scandinavia (Copenhagen, Gothenberg and Stockholm) and the USA (Boston, Chicago and New York) where cases were ascertained passively.[2,6,7] The characteristics of the populations, together with the weekly incidence of cases reported in these settings are summarized in Table 1 and Figure 1.

For the San Quentin prison, the first outbreak was reported in April 1918, with further outbreaks occurring in October and December 1918; the first influenza outbreaks in Scandinavian cities were reported in July 1918, with second waves occurring in October 1918; for settings in the US and for the Australian boats, the data relate to outbreaks occurring during October–November 1918. As shown in Table 1, the proportion of individuals reported to have experienced disease was high in community settings in Maryland (i.e. 22–59%) and on two of the boats (32% for the 'Medic' and 40% for the 'Boonah'). In contrast, this proportion was low (1–3%) in Stockholm and in cities in the USA, as reflected in the correspondingly lower weekly reporting rate shown in Figure 1. This low proportion may be partly attributable to the fact that cases were reported passively in cities in Scandinavia and USA (Table 1).

Estimates of the effective and basic reproduction numbers

The effective and basic reproduction numbers were first calculated for each study setting using the growth rate in the cumulative numbers of cases during the growth phase of the outbreak[8] (see below). The estimates obtained using this method are relatively insensitive to the level of under-reporting of cases, since the growth rate in the cumulative number of reported cases should be identical to that in the true cumulative number of cases in the population, unless the level of reporting changed over time. On the other hand, the estimates are sensitive to assumptions about the latent and infectious periods.[9]

To validate these estimates, we therefore recalculated the reproduction numbers using an approach which was independent of assumptions about the latent and infectious periods,

Table 1 Summary of the data used to estimate the basic reproduction number of Spanish influenza

Type of population	Study setting	Observation period	Percentage of the population experiencing disease (number of reported cases/population size)	Method of data collection
Confined	San Quentin prison	15/4/1918–27/5/1918	**5.2** (99/1900[a])	Passive: inmates reported to the prison infirmary when ill
		11/11/1918–15/12/1918	**3.7** (71/1900[a])	
		9/11/1918–4/12/1918	**3.1** (58/1900[a])	
	The 'Boonah' sailing boat	29/11/1918–8/1/1919	**39.5** (433/1095)	Passive and daily inspections
	The 'Devon' sailing boat	14/10/1918–31/10/1918	**7.4** (81/1096)	Passive and daily inspections
	The 'Medic' sailing boat	11/11/1918–15/12/1918	**31.5** (312/989)	Passive and daily inspections
Communities in Maryland	Baltimore	1/9/1918–30/11/1918	**22.2** (7489/33 776)	Active: house to house survey
	Cumberland		**39.8** (2085/5234)	
	Frederick		**31.7** (768/2 420)	
	Lonaconing		**59.4** (1093/1840)	
	Salisbury		**44.1** (765/1 735)	
Scandinavian cities	Copenhagen	30/6/1918–31/8/1918	**4.7** (25 147/539 000)	Routine notification data
		1/9/1918–28/12/1918	**13.3** (71 707/539 000)	
	Gothenberg	30/6/1918–31/8/1918	**2.3** (4657/198 948)	
		1/9/1918–28/12/1918	**9.8** (19 484/198 948)	
	Stockholm	30/6/1918–31/8/1918	**0.3** (1293/413 163)	
		1/9/1918–28/12/1918	**1.0** (4217/413 163)	
Cities in the USA	Boston	29/9/1918–2/11/1918	**0.4** (3327/767 813)	Routine notification data
		3/11/1918–25/1/1919	**1.6** (11 952/76 7813)	
	Chicago	15/9/1918–16/11/1918	**1.5** (37 907/2 596 681)	
		17/11/1918–3/4/1919	**0.7** (18 937/2 596 681)	
	New York	22/9/1918–23/11/1918	**2.0** (11 8453/5 872 667)	
		24/11/1918–29/03/1919	**0.5** (31 325/5 872 667)	

[a] New inmates entered the prison but their number and time of entry was not available

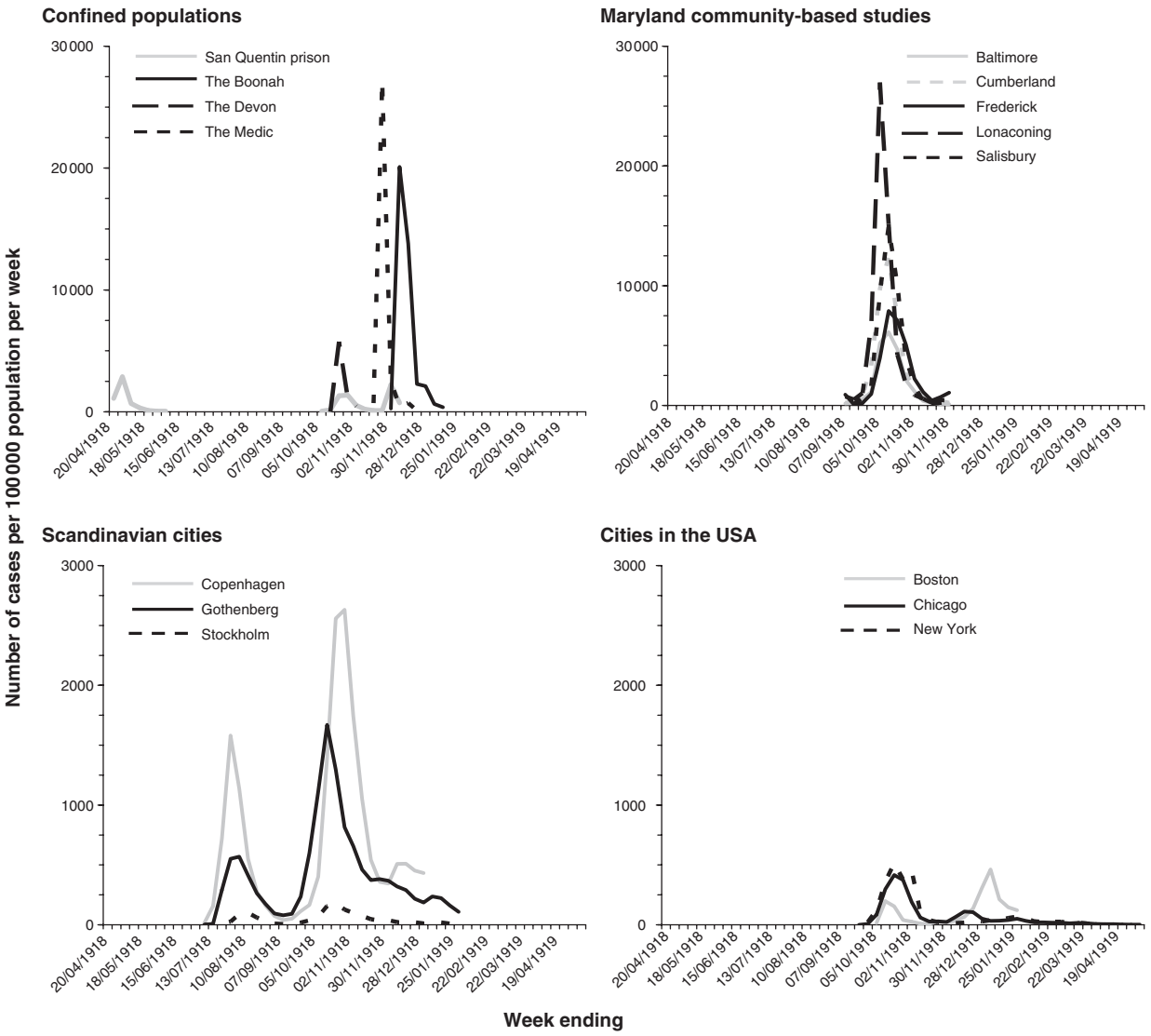

Figure 1 Summary of the weekly incidence of influenza cases observed in confined settings, in population-based studies in Maryland (USA), Scandinavian cities and in American cities during the period April 1918 to April 1919.

but which depended on reliable estimates of the final size of the outbreak,[10] for settings where there was likely to be little underreporting of cases (i.e. where at least 30% of the population was reported to have experienced influenza). For simplicity, the contribution of asymptomatic individuals is not considered here, as these affect estimates of the reproduction numbers based on the growth rate of the outbreak only if the proportion of individuals who are asymptomatic changes over time, and such changes are unlikely to have been substantial.

Estimates of the reproduction numbers using the growth rate of the epidemic

The effective reproduction number for influenza (R_n) at the start of each wave for the data shown in Figure 1 was calculated using the expression, as provided by Wearing *et al.*:[9]

$$R_n = \Lambda^2(L \times D) + \Lambda(L + D) + 1 \quad (1)$$

where L and D are the average durations of the latent and infectious periods, respectively and Λ is growth rate in the cumulative number of cases during the growth phase of the given outbreak. The above equation holds when the latent and infectious periods are assumed to follow the negative exponential distribution; in this situation, the term $L+D$ equals the serial interval, defined as the time interval between successive cases in a chain of transmission.[11,12] Equation (1) thus highlights the fact that the reproduction number increases with the size of the serial interval. The above expression for the effective reproduction number is analogous to that used to calculate the basic reproduction number using data from outbreaks for an infection which is introduced into a population for the first time, and has been applied to data for SARS and HIV.[13,14]

The average growth rate (Λ) in the cumulative numbers of cases in a given setting was calculated as the gradient of the straight line fitted, using linear regression, to the natural

logarithm of the, daily (where available) or weekly observed, cumulative number of cases, during the stage of the epidemic when the observed number of cases increased exponentially.

According to experimental studies, individuals are first infectious (i.e. shed virus) 1–2 days after inoculation (i.e. infection) and shed virus for 3 or 4 days thereafter.[15–23] We therefore assume that the average latent period is 2 days and, assuming that the infectious period follows the negative exponential distribution, the average infectious period is 2 days. The latter assumption is consistent with 61, 37, 22 and 14% of individuals still being infectious 1, 2, 3 and 4 days, respectively after onset of infectiousness. These assumptions for the latent and infectious periods imply an average serial interval of 4 days.

Given recent estimates of an average serial interval lasting 2.5 days[24] and the possibility that some individuals may shed virus for time periods which are longer than the average, we explore the sensitivity of our reproduction number estimates to the assumption that the average serial interval was 2.5 or 6 days, using an average infectious period of 0.5 and 4 days, respectively and an average latent period of 2 days.

Estimates of the basic reproduction number for each epidemic wave were then obtained using the expression:

$$R_0 = \frac{R_n}{\text{proportion susceptible at the start of the wave}}$$

It was assumed that 100% of individuals were susceptible at the start of April 1918. The extent to which this proportion declined thereafter, as a result of the increased incidence of influenza in some settings during the spring of 1918[2,25] is unknown. We here assume that 70% and 50% were susceptible at the start of the first and second subsequent waves after April 1918, which is consistent with the level of seroconversion found as a result of the first wave of the Hong Kong influenza pandemic[26] and the observations that, during a typical influenza A season, 20–30% of individuals experience rises in their titres of antibodies attributable to the circulating strain of influenza.[27] Other analyses of the R_0 for Spanish influenza[3] also used 70% as a lower bound for the proportion of the population which was susceptible at the start of outbreaks during the autumn of 1918.

The assumption that the latent and infectious periods follow the negative exponential distribution may lead to under- or over-estimates of the reproduction numbers, as calculated using equation (1).[9] To test the sensitivity of our reproduction number estimates to this assumption, we recalculated them using the following formula provided by Wearing *et al.*:[9]

$$R_n = \frac{\Lambda D\left\{(\Lambda L/m)+1\right\}^m}{[1-\{(\Lambda D/n)+1\}^{-n}]}$$

using the extreme assumptions $m=n=100$. This formula is related to the above formula for the net reproduction number, but allows the distributions of the latent and infectious periods to vary according to the size of m and n, respectively. The value $m=1$ (or $n=1$) corresponds to the assumption that the latent (or infectious period) follows the negative exponential distribution; the value $m=100$ (or $n=100$) corresponds to the assumption that the latent (or infectious) period is tightly distributed around the mean (Figure A1).

Estimates of the reproduction numbers using the final size of the epidemic

For settings with a high level of reporting, i.e. where >30% of the population was reported to have experienced influenza [the 'Boonah', the 'Medic' and all the Maryland study sites, apart from Baltimore (Table 1)], the R_0 was calculated using the final epidemic size using the approach of Becker:[10]

$$R_0 = \frac{N-1}{C}\sum_{i=S_f+1}^{S_0} 1/i \qquad (2)$$

where N is the population size, C is the total number of cases in the given wave and S_0 and S_f are the numbers of susceptible individuals in the population at the start and end of the given wave, respectively.[10] The standard error in the R_0 was calculated using the expression:[10]

$$SE(R_0) = \frac{N-1}{C}\sqrt{\sum_{i=S_f+1}^{S_0} 1/i^2 + \frac{CR_0^2}{(N-1)^2}} \qquad (3)$$

The effective reproduction number and its standard error were calculated using expressions (2) and (3), substituting the term for the number of susceptible individuals at the start of the outbreak with the total population size.

Results

Figure 2A summarizes estimates of the effective reproduction number for the settings considered here.

In general, the estimates were low for community-based settings, ranging between 1.2 and 3.0, based on the assumption that the average serial interval was 4 days. For confined settings and for the same assumption about the serial interval, the effective reproduction number ranged between 2.1 and 7.5, with the highest value estimated for ships (e.g. 3.8 for 'The Medic' and 7.5 for 'The Devon'). Excepting that estimated for the San Quentin prison, the effective reproduction number associated with the second wave of the pandemic in each setting was lower than that for the first (e.g. 3.0 in Gothenberg in July 1818 vs 1.5 in October 1918), which is attributable to the correspondingly lower proportion of individuals who are susceptible to infection at the start of the second wave.

The implications of these estimates for the basic reproduction number, based on the assumption that 100, 70 and 50% of the population was susceptible at the start of April 1918 and the first and second subsequent waves, respectively, are summarized in Figure 2B. For all community-based settings, the R_0, as calculated using the growth rate of the outbreak, was between 2 and 4.3, assuming that the serial interval was either 2.5 or 4 days. For the assumption that the serial interval was 6 days, the lowest basic reproduction number (≈2.4) was estimated for Baltimore, Cumberland, Frederick and for New York in October 1918 and the highest (4.3) was estimated for Gothenberg in July 1918.

For confined settings, the basic reproduction number, as calculated using the growth rate of the given wave, was lowest for the San Quentin prison in April 1918 and highest for the outbreak on 'The Devon', occurring in the ranges 1.9–5.6

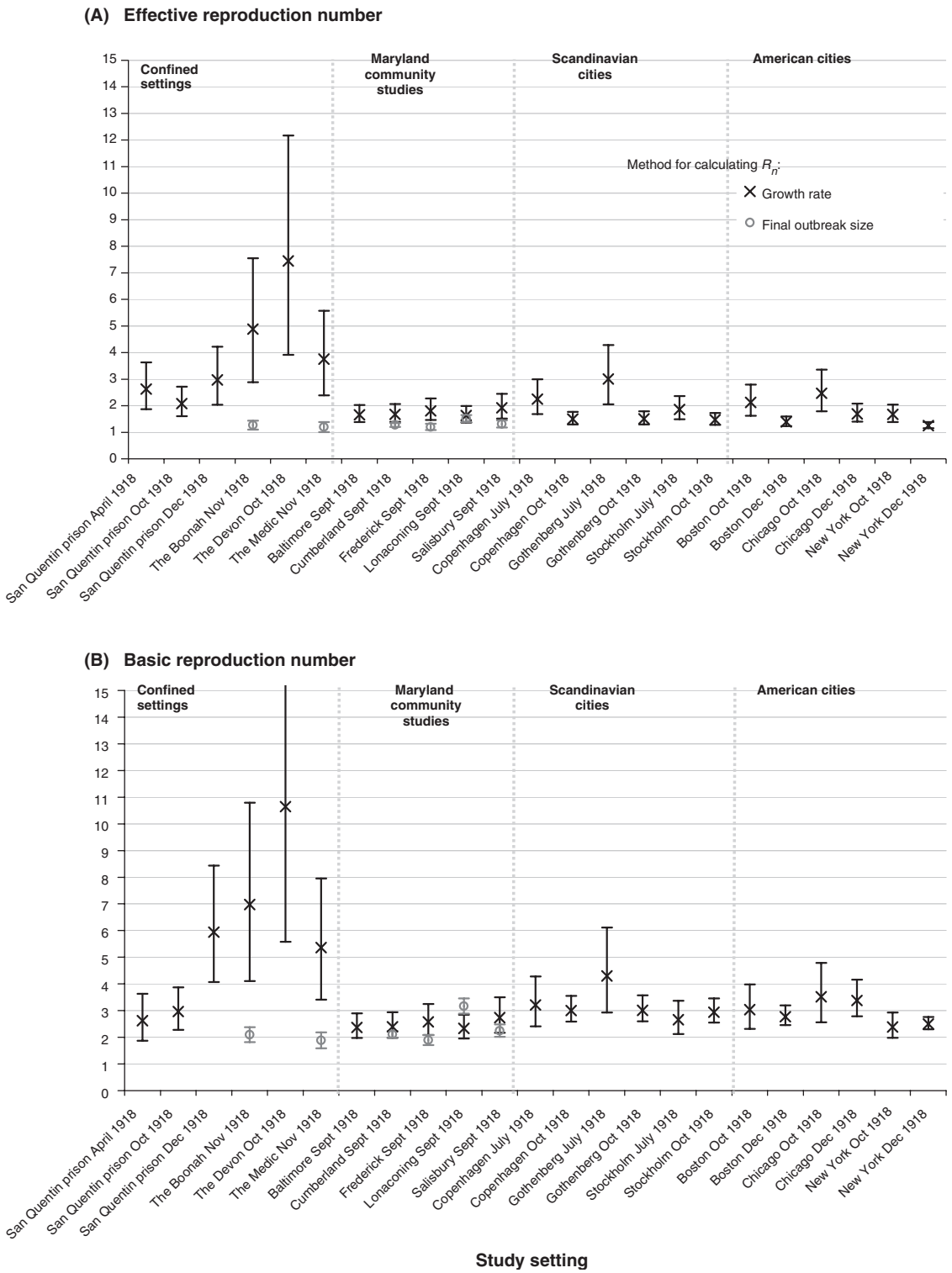

Figure 2 Summary of estimates of the effective and basic reproduction number in various settings during the period April 1918 to April 1919. The lower and upper limits on the points for estimates obtained using the growth rate in the cumulative numbers of cases reflect reproduction numbers obtained assuming a serial interval of 2.5 and 6 days, respectively. For the 'Devon', the upper limit of the basic reproduction number estimate is 17.

and 2.6–10.6 for the assumption that the serial interval was 2.5 and 4 days, respectively. The assumption of a 6 day serial interval was associated with very high R_0 estimates for all outbreaks in confined settings (e.g. >15 for 'The Devon').

For community-based settings, estimates of the reproduction numbers calculated using the final size of the outbreak were consistent with those estimated using the growth rate of the wave (Figure 2). For the outbreaks occurring on the boats,

on the other hand, they were substantially lower than those estimated using the growth rate of the outbreak (e.g. an R_0 of 2.1 vs 7.0 for 'The Boonah', as calculated using the outbreak size and the growth rate of the outbreak, respectively assuming an average serial interval of 4 days).

As shown in Figure A2 in the appendix, estimates for the reproduction numbers were relatively insensitive to assumptions about the distribution of the latent and infectious periods, for all settings except for 'The Devon'. In this setting, the assumption that the latent and infectious periods were tightly distributed resulted in an R_0 estimate of 16.7, as compared with 10.6 if it was assumed that they followed the negative exponential distribution.

Discussion

Our analyses illustrate the application of different methods for estimating the reproduction numbers of Spanish influenza, which are based on the growth rate and the final size of the epidemic. These methods are slightly simpler than those used in previous estimates of the reproduction number for Spanish influenza,[3–5] which involved fitting predictions from a transmission model to the observed data; those based on the growth rate of the epidemic also had the advantage of being relatively insensitive to the level of underreporting of cases, provided that this did not change over time. Our estimates are of the same order as the estimates from those studies, confirming the conclusion that, in the general population, despite its high pathogenicity, the transmissibility of the Spanish influenza virus was relatively low, with a basic reproduction number lying in the range 2–4.

The reproduction numbers were estimated for all settings using the growth rate in the cumulative numbers of cases during the growth phase of each pandemic wave. The advantage of this method is that the estimates were relatively insensitive to the level of underreporting, since the growth rate in the cumulative number of reported cases should be identical to that in the true cumulative number of cases in the population, unless the level of underreporting changed over time. Changes in the reporting of cases are likely to have been small, at least in the city-based community studies, where cases were ascertained passively, given our finding that the reproduction numbers in these settings were similar to those in Maryland, where cases were ascertained actively through house-to-house surveys and where the level of underreporting was probably low.[7]

In our analyses of the reproduction numbers based on the growth rate in the cumulative numbers of cases, the average serial interval was assumed to be 4 days. This assumption is consistent with data presented in three family-based studies during the Asian and Hong Kong influenza pandemics,[28–30] for which, considering cases who had onset within 10 days of the first case in the household, the median time interval between the first and subsequent cases was approximately 4 days. Our finding, that at least for community-based studies in Maryland, where the reporting of cases was likely to have been reliable, the reproduction numbers, as calculated using the outbreak size (and independently of the serial interval), were consistent with those calculated using the growth rate of the cumulative numbers of cases (Figure 2), provides further justification for our assumed values for the serial interval. We discuss the discrepancy between the estimates obtained using the two methods for confined settings below.

In practice, the serial interval depends on many factors, e.g. the time when symptoms are their most severe and individuals isolate themselves from others and the amount of viral shedding, which appears to be correlated with the severity of the illness[17,21] and varies throughout the period of viral shedding.[17–22] It is also possible that the duration of infectiousness is shorter for epidemic than for pandemic influenza, given that virus replication and shedding may be limited if individuals have some protection resulting from previous exposure to a strain, which is antigenically related to that causing the current infection. This may partly explain the recent low estimates of the serial interval of 2.5 days for influenza,[24] which were based on household data of epidemic influenza transmission from France in the year 2002.

Our estimates of the reproduction numbers for the boats based on the growth rate of the outbreak were much higher than those estimated for population-based settings (Figure 2). This is consistent with the crowded conditions on these boats, and hence the increased opportunity for exposure, as reflected in the high attack rates.

The reproduction numbers calculated for the boats using the growth rate of the outbreak (where the latent and infectious periods were parameters) were substantially higher than those calculated using the final outbreak size. The discrepancy could be due to genuine reductions in the (basic) reproduction number during the outbreak, which would have affected the estimates calculated using the final outbreak size more than those calculated using the growth rate in the cumulative numbers of cases. Such reductions could have occurred because of reduced contact, once individuals became aware of the outbreak, or control measures. For example, daily 'thermometer parades' were carried out on each of the boats, and any individuals with signs of fever or reporting sick were removed for observation or immediately isolated.[2]

Overall, the analyses presented here support the findings from other studies, which were based on alternative data sources and methods, that the R_0 for Spanish influenza was relatively low when the virus first appeared in the general population, ranging between 2 and 4.3. Many factors have changed since 1918, e.g. household sizes have probably decreased over time, although the spread of the influenza virus may now be rapid because of the increased connectivity of individuals. Control measures which are now available are, in theory, more effective than those available in 1918, which consisted of shutting down music halls and theatres and staggering the opening times of offices and shops to reduce aggregation of individuals.[2] It is hoped that good planning, basic infection control measures and the availability of antivirals will help to ensure that any future influenza pandemic has a smaller impact than that of the Spanish influenza pandemic.

Acknowledgements

We thank Dr WJ Edmunds and Dr B Cooper (Health Protection Agency Centre for Infections) for helpful comments on the manuscript.

Conflict of interest: None declared.

KEY MESSAGES

- This study uses morbidity data from cities in Europe and America and from confined settings, and several different methods based on the growth rate and final size of the epidemic, to estimate the effective and basic reproduction numbers for the 1918 (Spanish) influenza virus.
- The effective reproduction number (the average number of secondary infectious cases produced by a typical infectious case in a given population) for the 1918 influenza virus was in the range 1.2–3.0 and 2.1–7.5 for community-based and confined settings, respectively.
- The basic reproduction number (the average number of secondary infectious cases resulting from a typical infectious case in a totally susceptible population) was in the range 2.4–4.3 and 2.6–10.6 cases in community-based and confined settings, respectively.
- These findings for community-based populations confirm the relatively low transmissibility of the 1918 (Spanish) influenza virus, which has been found by other studies using alternative data sources and methods.

References

[1] Osterholm MT. Preparing for the next pandemic. *N Engl J Med* 2005;**352:**1839–42.

[2] Ministry of health. *Report on the Pandemic of Influenza*. London: His Majesty's Stationery Office, 1920, Report no. 4.

[3] Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza. *Nature* 2004;**432:**904–6.

[4] Spicer CC, Lawrence CJ. Epidemic influenza in Greater London. *J Hyg (Lond)* 1984;**93:**105–12.

[5] Chowell G, Ammon C, Hengartner N, Hyman J. Estimation of the reproductive number of the Spanish flu epidemic in Geneva, Switzerland. *Vaccine* 2006;**24:**6747–50.

[6] Stanley L. Influenza at San Quentin prison, California. *Public Health Rep* 1919;**34:**996–1008.

[7] Frost W, Sydenstricker E. Influenza in Maryland. Preliminary statistics of certain localities. *Public Health Rep* 1919;**34:**491–504.

[8] Dietz K. The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res* 1993;**2:**23–41.

[9] Wearing HJ, Rohani P, Keeling MJ. Appropriate models for the management of infectious diseases. *PLoS Med* 2005;**2:**621–27.

[10] Becker N. *Analysis of Infectious Disease Data*. London: Chapman and Hall, 1989.

[11] Hope-Simpson RE. The period of communicability of certain epidemic diseases. *Lancet* 1948;**13:**755–60.

[12] Fine PE. The interval between successive cases of an infectious disease. *Am J Epidemiol* 2003;**158:**1039–47.

[13] Anderson RM, Medley GF, May RM, Johnson AM. A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA J Math Appl Med Biol* 1986;**3:**229–63.

[14] Lipsitch M, Cohen T, Cooper B *et al*. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003;**300:**1966–70.

[15] Murphy BR, Baron S, Chalhub EG, Uhlendorf CP, Chanock RM. Temperature-sensitive mutants of influenza virus. IV. Induction of interferon in the nasopharynx by wild-type and a temperature-sensitive recombinant virus. *J Infect Dis* 1973;**128:**488–93.

[16] Alford RH, Kasel JA, Gerone PJ, Knight V. Human influenza resulting from aerosol inhalation. *Proc Soc Exp Biol Med* 1966;**122:**800–4.

[17] Couch RB, Douglas RG Jr, Fedson DS, Kasel JA. Correlated studies of a recombinant influenza-virus vaccine. 3. Protection against experimental influenza in man. *J Infect Dis* 1971;**124:**473–80.

[18] Calfee DP, Peng AW, Cass LM, Lobo M, Hayden FG. Safety and efficacy of intravenous zanamivir in preventing experimental human influenza A virus infection. *Antimicrob Agents Chemother* 1999;**43:**1616–20.

[19] Hayden FG, Treanor JJ, Betts RF, Lobo M, Esinhart JD, Hussey EK. Safety and efficacy of the neuraminidase inhibitor GG167 in experimental human influenza. *JAMA* 1996;**275:**295–99.

[20] Hayden FG, Treanor JJ, Fritz RS *et al*. Use of the oral neuraminidase inhibitor oseltamivir in experimental human influenza: randomized controlled trials for prevention and treatment. *JAMA* 1999;**282**(13):1240–46.

[21] Hayden F, Fritz R, Lobo M, Alvord W, Strober W, Straus SE. Local and systemic cytokine responses during experimental human influenza A virus infection. *J Clin Invest* 1998;**101:**643–49.

[22] Treanor JJ, Hayden FG, Vrooman PS *et al*. Efficacy and safety of the oral neuraminidase inhibitor oseltamivir in treating acute influenza: a randomized controlled trial. US Oral Neuraminidase Study Group. *JAMA* 2000;**283:**1016–24.

[23] Reeve P, Gerendas B, Moritz A *et al*. Studies in man with cold-recombinant influenza virus (H1N1) live vaccines. *J Med Virol* 1980;**6:**75–83.

[24] Ferguson NM, Cummings DA, Cauchemez S *et al*. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 2005;**437:**209–14.

[25] Olson DR, Simonsen L, Edelson PJ, Morse SS. Epidemiological evidence of an early wave of the 1918 influenza pandemic in New York City. *Proc Natl Acad Sci USA* 2005;**102:**11059–63.

[26] Stuart-Harris C. Epidemiology of influenza in man. *Br Med Bull* 1979;**35:**3–8.

[27] Monto AS, Koopman JS, Longini IM Jr. Tecumseh study of illness. XIII. Influenza infection and disease, 1976–1981. *Am J Epidemiol* 1985;**121:**811–22.

[28] Jordan WS Jr, Denny FW Jr, Badger GF *et al*. A study of illness in a group of Cleveland families. XVII. The occurrence of Asian influenza. *Am J Hyg* 1958;**68:**190–212.

[29] Davis LE, Caldwell GG, Lynch RE, Bailey RE, Chin TD. Hong Kong influenza: the epidemiologic features of a high school family study analyzed and compared with a similar study during the 1957 Asian influenza epidemic. *Am J Epidemiol* 1970;**92:**240–47.

[30] Woodall J, Rowson KEK, McDonald J. Age and Asian influenza, 1957. *Br Med J* 1958;**5108:**1316–18.

Appendix

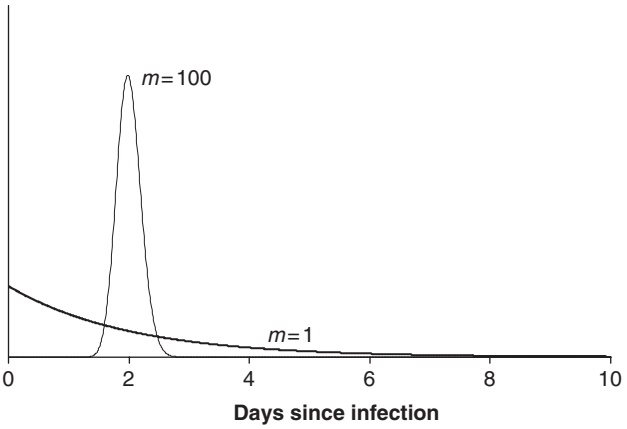

Figure A1 Comparison between the distributions of the latent period obtained assuming that the values for $m = 1$ and 100 and the average latent period is 2 days. The formula [9] for this distribution is $((m/L)^m\ e^{-mt/L}\ t^{m-1})/(m-1)!$.

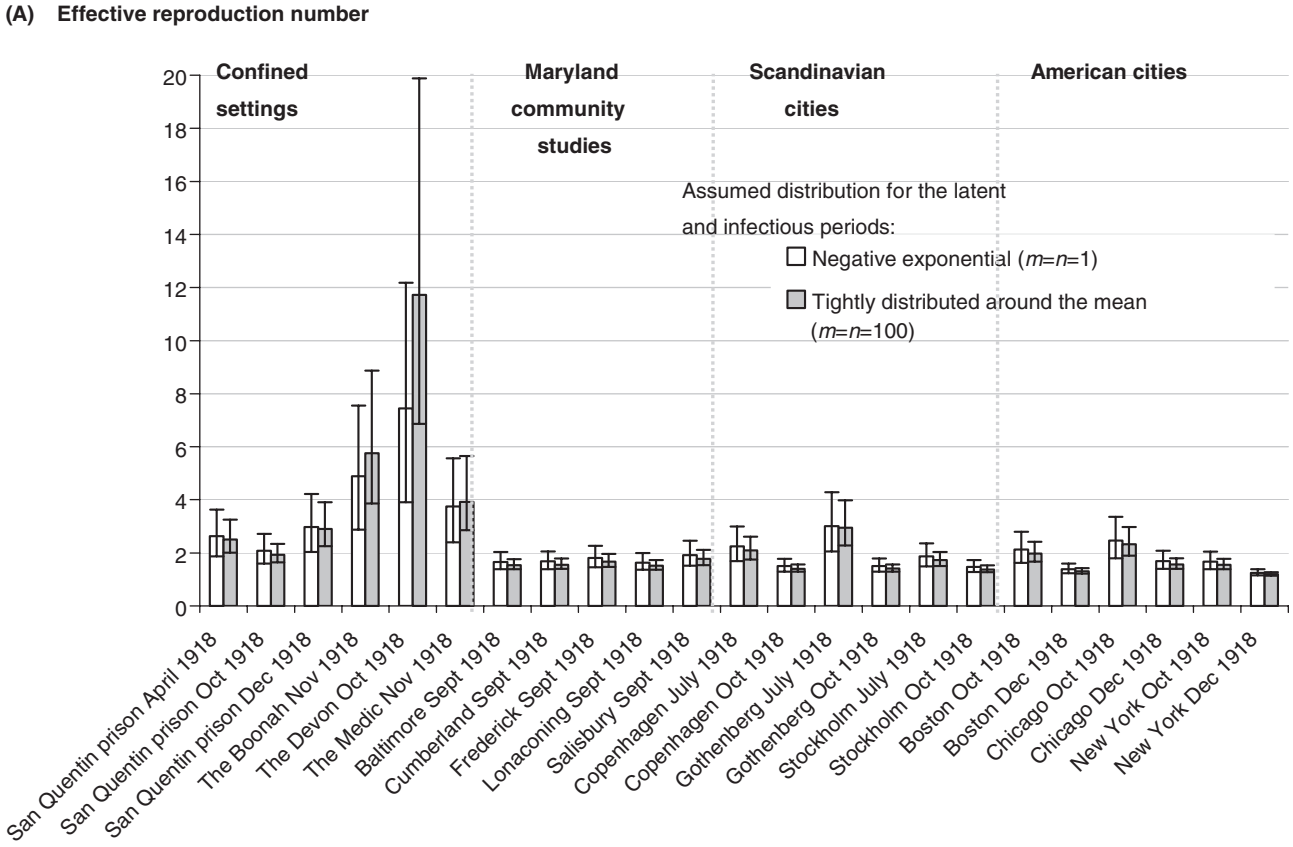

Figure A2 Comparison between the effective and basic reproduction numbers as calculated using the growth rate of the epidemic wave, assuming that the latent and infectious periods either followed the negative exponential distribution or were tightly distributed around the mean value ($m = n = 100$). The lower and upper limits for the estimates reflect reproduction numbers obtained assuming a serial interval of 2.5 and 6 days, respectively.

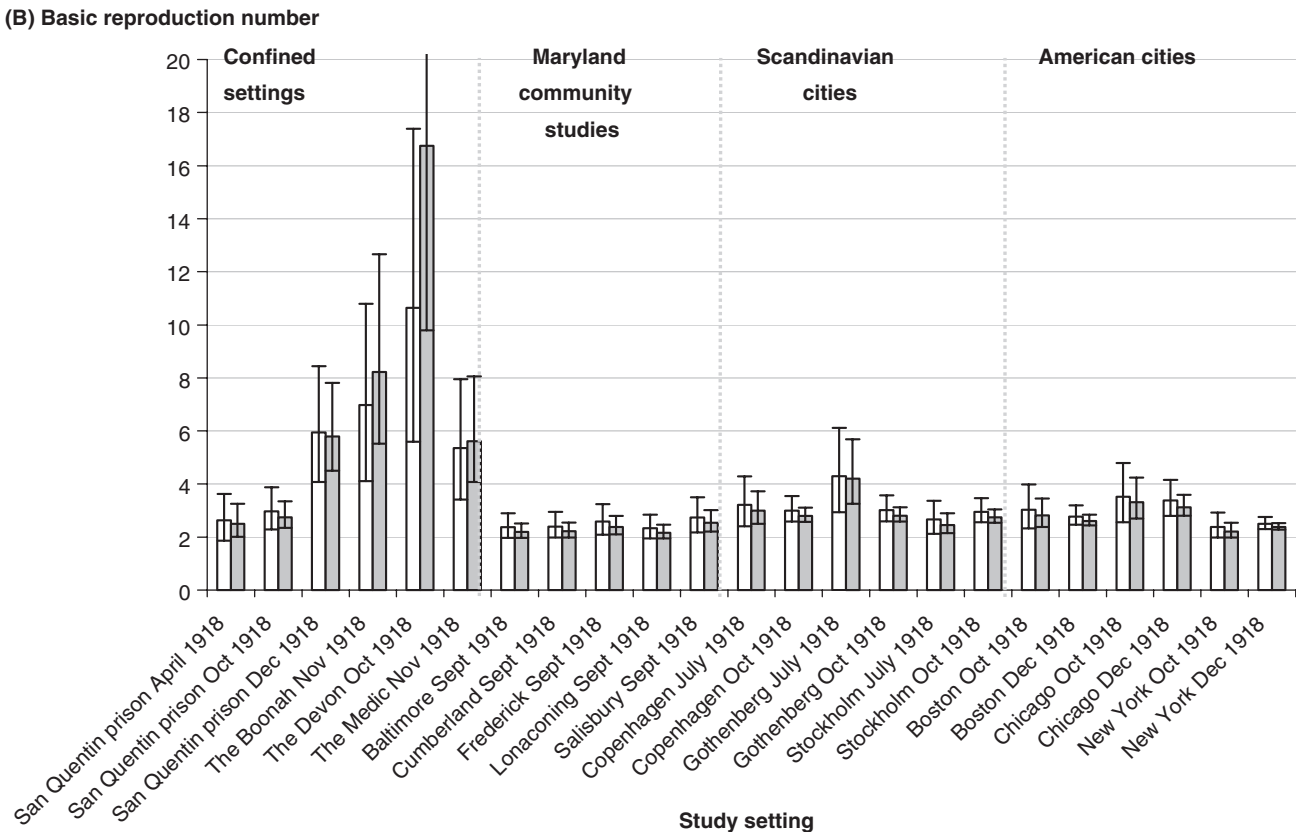

Figure A2 Continued.