PROCEEDINGS OF THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact

Gerardo Chowell, Luís M.A Bettencourt, Niall Johnson, Wladimir J Alonso and Cécile Viboud

Supplementary data	"Data supplement" http://rspb.royalsocietypublishing.org/content/suppl/2009/03/04/275.1634.501.DC1.html
References	**This article cites 32 articles, 9 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/275/1634/501.full.html#ref-list-1 **Article cited in:** http://rspb.royalsocietypublishing.org/content/275/1634/501.full.html#related-urls
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact

Gerardo Chowell[1,2,*], Luís M. A. Bettencourt[2,3], Niall Johnson[4], Wladimir J. Alonso[5] and Cécile Viboud[5]

[1]*School of Human Evolution and Social Change, Arizona State University, PO Box 872402, Tempe, AZ 85287, USA*
[2]*Center for Nonlinear Studies and Mathematical Modeling and Analysis (MS B284), Theoretical Division,*
Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[3]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*
[4]*Australian Commission on Safety and Quality in Health Care, GPO Box 5480, Sydney, New South Wales 2001, Australia*
[5]*Fogarty International Center, National Institutes of Health, 16 Center Drive, Bethesda, MD 20892, USA*

Spatial variations in disease patterns of the 1918–1919 influenza pandemic remain poorly studied. We explored the association between influenza death rates, transmissibility and several geographical and demographic indicators for the autumn and winter waves of the 1918–1919 pandemic in cities, towns and rural areas of England and Wales. Average measures of transmissibility, estimated by the reproduction number, ranged between 1.3 and 1.9, depending on model assumptions and pandemic wave and showed little spatial variation. Death rates varied markedly with urbanization, with 30–40% higher rates in cities and towns compared with rural areas. In addition, death rates varied with population size across rural settings, where low population areas fared worse. By contrast, we found no association between transmissibility, death rates and indicators of population density and residential crowding. Further studies of the geographical mortality patterns associated with the 1918–1919 influenza pandemic may be useful for pandemic planning.

Keywords: influenza; Spanish flu pandemic; reproduction number; demographics; scaling laws;
England and Wales

1. INTRODUCTION

Greater understanding of past influenza pandemic patterns is key for designing more efficient public health interventions for future outbreaks. Despite the growing interest in historical epidemiological studies, spatial variations in the transmissibility and mortality rates of the 1918–1919 influenza pandemic have yet to be carefully quantified. In addition, little is known about the comparative dynamics of influenza in cities and rural areas in general, mostly due to the lack of appropriate data at this scale for inter-pandemic and pandemic seasons.

Transmissibility can be quantified by the reproduction number (R), the average number of secondary cases generated by an index case, which determines the intensity and types of interventions required to avert an epidemic (Anderson & May 1991; Diekmann & Heesterbeek 2000). Past research has estimated that the reproduction number of the 1918–1919 influenza pandemic ranged between 1.5 and 5.4 (Mills *et al.* 2004; Gani *et al.* 2005; Chowell *et al.* 2006, 2007*a*; Ferguson *et al.* 2006; Sertsou *et al.* 2006; Viboud *et al.* 2006*a*; Massad *et al.* 2007; Nishiura 2007; Andreasen *et al.* in press), depending on the specific location and pandemic wave considered, type of data,

estimation method and level of spatial aggregation, which has ranged from small towns of a few hundred people to entire nations with several million inhabitants. The variability in published R estimates suggests that local factors, including geographical and demographic characteristics, could play a role in disease transmissibility.

In parallel, large variations in the 1918–1919 pandemic mortality rates have been reported between different nations and cities in the US, and were linked to differences in socio-demographic conditions and public health interventions (Murray *et al.* 2006; Bootsma & Ferguson 2007; Hatchett *et al.* 2007). The differences between urban and rural areas and population units of varying size, however, have not been comprehensively analysed.

Here we quantify geographical patterns during the autumn and winter waves of the 1918–1919 pandemic in English and Welsh cities, towns and rural areas. We analyse weekly influenza mortality from a diversity of geographical units to explore how transmissibility and mortality rates varied locally with socio-demographic factors, including population size, population density, residential crowding and urbanization.

2. MATERIAL AND METHODS

(a) *Data sources*

National vital registration began in 1837 in England and Wales. Weekly numbers of influenza-specific deaths covering

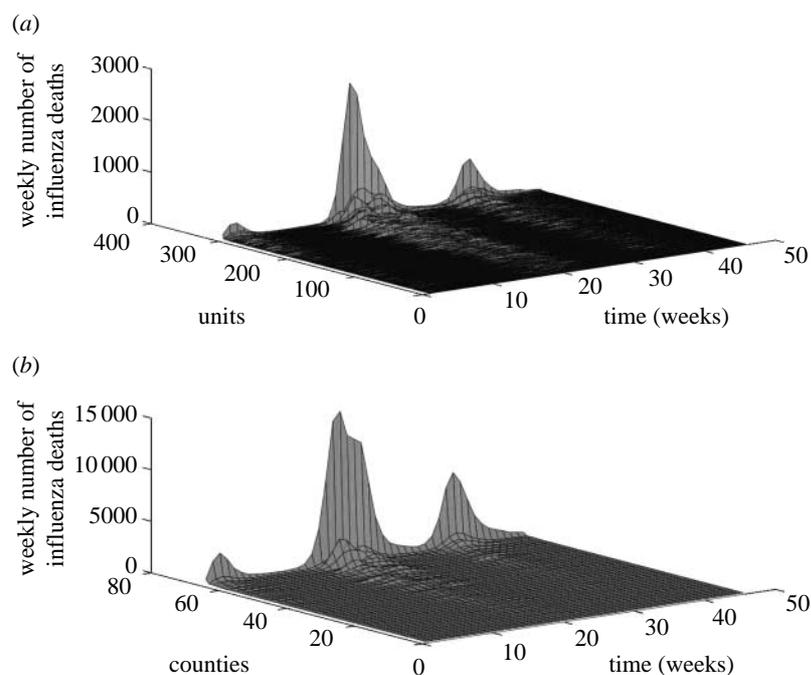* Author for correspondence (gchowell82@asu.edu).

Figure 1. Weekly number of influenza-specific deaths during the 1918–1919 influenza pandemic in England and Wales at two different spatial scales: (*a*) 305 administrative units and (*b*) 62 counties (Ministry of Health 1920). The influenza mortality data cover 46 weeks comprising the week ending 29 June 1918 and continuing to the week ending 10 May 1919.

46 weeks from 29 June 1918 to 10 May 1919 were compiled from an official publication of the British Ministry of Health (Ministry of Health 1920), and have been made partially available online (Influenza Pandemic Mortality in England and Wales 2006, http://ahds.ac.uk/catalogue/collection.htm?uri=hist-4350-11918-1919). In England and Wales, as in other European countries, the 1918–1919 pandemic occurred in three major waves (Low 1920; Andreasen *et al.* in press; figure 1): a summer wave, associated with low mortality (29 June to 3 August 1918); an autumn wave with high mortality (12 October to 28 December 1918); and a winter wave with intermediate mortality (8 February to 5 April 1919). Here we focus on the autumn 1918 and winter 1919 waves due to better resolution in the mortality data.

The mortality data covered all of England and Wales, and were available at the refined spatial scale of administrative units ($N=305$), as well as the coarser scale of counties ($N=62$) (Ministry of Health 1920). A total of 247 towns and cities with more than 10 000 inhabitants were categorized as urban units, and London was considered as a single unit. The remaining 58 units comprised rural areas, including 42 'remainder of counties' classified as non-urban locales by the Ministry of Health and 16 peripheral administrative counties that did not contain a single city or town (of which 11 were located in Wales).

Demographic data for the 305 administrative units and 62 counties were retrieved from the 1921 decennial census (A Vision of Britain 2006), and included total population size, surface area, number of dwellings and number of rooms occupied. From these data, we derived measures of population density and residential crowding (table 1) and calculated the weekly and cumulative death rates for each geographical unit and pandemic wave. Although no census was conducted in 1918, the demographic data from the 1911 and 1921 census had very similar geographical patterns (correlation > 0.99), so that our 1921 demographic estimates were deemed appropriate for the study of the 1918–1919

pandemic geographical patterns. Given the spatial aggregation of the dataset, some of the rural units had large total population sizes, but population density (defined as the ratio of population size to surface area) was 14 times lower on average in rural areas than in cities and towns.

(b) Transmissibility of the autumn and winter pandemic waves (reproduction number, R)

(i) *Estimation methods*

We estimated the reproduction number in each geographical unit and during each pandemic wave, relying on the early ascending phase of local epidemic curves, where saturation effects arising from the depletion of susceptible individuals can be neglected (Anderson & May 1991). We assumed an initial exponential growth and estimated the intrinsic growth rate, '*r*', by fitting a straight line (with slope *r* and intercept *b*) to the initial increase in weekly deaths (in logarithmic scale). The longest epidemic period consistent with exponential growth was determined by the goodness-of-fit test statistic (Chowell *et al.* 2007*a*). The reproduction number was calculated by substituting the estimate for *r* into an expression derived from the linearization of the classical susceptible–exposed–infectious–recovered (SEIR) transmission model (Lipsitch *et al.* 2003; Wallinga & Lipsitch 2007)

$$R = \left(1 + \frac{r}{b_1}\right)\left(1 + \frac{r}{b_2}\right), \tag{2.1}$$

where $1/b_1$ and $1/b_2$ are the mean latent and infectious periods, respectively. This expression for R assumes exponentially distributed latent and infectious periods, and the mean generation interval between two successive cases is given by $T_c = 1/b_1 + 1/b_2$. We carried out a sensitivity analysis to assess the impact of this assumption on the R estimates. We derived an upper bound for the extreme case of a fixed generation interval (delta distribution) using the following equation (Wallinga & Lipsitch 2007):

$$R = e^{rT_c}. \tag{2.2}$$

Table 1. The range, median and interquartiles of demographic variables in 305 administrative units of England and Wales used in our analyses. (Data from A Vision of Britain 2006).

	range			interquartile range	
variable	min	max	median	lower bound	upper bound
population size	11 427	4 484 523	52 710	31 158	125 517
population density (people/statute acres)	0.08	64.25	10.46	3.62	17.53
average number of people per dwelling	2.03	9.01	4.79	4.38	5.14
average number of people per room	0.71	1.97	0.94	0.83	1.06

There is considerable debate about the duration of the generation interval for influenza (Andreasen *et al.* in press), with the most recent estimates relying on seasonal influenza transmission in households (Cauchemez *et al.* 2004; Ferguson *et al.* 2006), while no data are available for pandemic influenza. Given the uncertainty, we considered two extreme values of the generation interval used in past research: a short interval of 3 days (where the latent and infectious periods were both set to 1.5 days; Ferguson *et al.* 2005, 2006; Wallinga & Lipsitch 2007) and a longer interval of 6 days (latent period = 1.9 days and infectious period = 4.1 days; Longini *et al.* 2004; Mills *et al.* 2004). The case fatality rate was set to 2%, as in a previous transmissibility study (Mills *et al.* 2004).

(ii) *Heterogeneity in influenza transmissibility and association with socio-demographic factors*
We explored whether there was a meaningful variability in the reproduction number of the 1918–1919 influenza pandemic across England and Wales by comparing the variability in the *R* estimates between and within counties. We estimated the within-county variability from the finer spatial scale of administrative units, and used the analysis of variance (ANOVA) to test the county-specific differences in transmissibility (Neter & Wasserman 1974).

The association between reproduction number and socio-demographic variables was explored via Spearman correlations, using a Bonferroni correction for multiple comparisons.

(c) *Cumulative influenza death rates for the autumn and winter pandemic waves*
(i) *Death rates, population size and urbanization*
We initially explored the association between death rates, census variables and urbanization, by correlation and multivariate regression. We identified population size and urbanization as statistical predictors of death rates. We characterized these relationships further by applying two methods derived from econometrics and previously applied in infectious disease epidemiology, the Lorenz curve and the summary Gini index (Lee 1997; Woolhouse *et al.* 1997; Kerani *et al.* 2005; Green *et al.* 2006). The Lorenz curve is a graphical representation of the cumulative distribution of a quantity representing in our case the proportion of deaths assumed by the bottom y% of population sizes. If all death rates are perfectly equal, the Lorenz curve is the first diagonal (no heterogeneity), while a perfectly unbalanced distribution appears as a vertical line (maximum heterogeneity, for instance, if all units except one had zero deaths). Most empirical distributions of death rates lie somewhere in between.

The Gini index is a summary statistic of the Lorenz curve, ranging between 0 and 1, calculated as the area between the Lorenz curve and the diagonal representing perfect equality.

A large Gini index indicates highly heterogeneous death rates. Conversely, a Gini index of zero represents an absence of heterogeneity, so that death rates are directly proportional to population size.

(ii) *Scaling laws*
Next, we characterized the functional relationship between population size and cumulative number of influenza deaths. Influenza deaths follow a power-law function of population sizes if $D \propto N^g$, where D denotes number of deaths; N indicates population size; and g is an exponent to be estimated. For $g = 1.0$, deaths are exactly proportional to population size (i.e. no heterogeneity); $g < 1$ indicates that low population units have higher per capita death rates, while the opposite is true for $g > 1$.

(d) *Timing of pandemic onset and socio-demographic factors*
Finally, for the autumn and winter waves, we explored whether the timing of the local pandemic onset varied with socio-demographic characteristics. The onset week was defined for each pandemic wave and geographical unit as the first week associated with a monotonic increase in weekly deaths, up to the week of peak deaths.

3. RESULTS
(a) *Transmissibility of the autumn and winter pandemic waves*
(i) *Reproduction number estimates* (R)
Figure S1 in the electronic supplementary material shows the distribution of the reproduction number estimates for the autumn and winter waves of the 1918–1919 pandemic in England and Wales, at the refined spatial scale of administrative units ($N = 305$) and the coarser scale of counties ($N = 62$). Most locations experienced an initial epidemic phase lasting at least three epidemic weeks, so that R estimates derived from equation (2.1) could be obtained in 87% of administrative units and 100% of the counties for the autumn wave, and 69% of the units and 87% of the counties for the winter wave. Normal probability plots indicated that R estimates at the county level closely followed a normal distribution, while the distributions of estimates at the administrative unit level showed a greater frequency of high values compared with a normal distribution (not shown). In table 2, the summary estimates of R are presented for a short and long duration of the generation interval (3 and 6 days). For the shorter generation interval in the autumn wave, the mean R was found to be 1.40 (95% CI: 1.38–1.42) in the administrative units, with similar values at the county level. The mean R estimate based on the aggregated national pandemic wave was not different, at 1.39

Table 2. Reproduction number (*R*) estimates for the autumn and winter 1918–1919 pandemic waves with 95% CIs for England and Wales, at three different spatial scales and using different assumptions about the generation interval.

	3-day generation interval			6-day generation interval		
	administrative units ($N^a=305$)	counties ($N^a=62$)	national ($N^a=1$)	administrative units ($N^a=305$)	counties ($N^a=62$)	national ($N^a=1$)
autumn wave	1.40 (1.38, 1.42)	1.36 (1.33, 1.39)	1.39 (1.36, 1.43)	1.86 (1.82, 1.90)	1.76 (1.69, 1.83)	1.84 (1.75, 1.92)
winter wave	1.35 (1.33, 1.37)	1.33 (1.31, 1.36)	1.39 (1.29, 1.49)	1.74 (1.70, 1.78)	1.71 (1.66, 1.76)	1.82 (1.61, 2.05)

[a] *R* was estimated in 87% (264/305) of the administrative units and 100% of the counties (62/62) during the autumn wave, and in 69% (209/305) and 87% (54/62) of the administrative units and the counties during the winter wave, respectively.

Table 3. Relationship between influenza transmissibility (*R* estimates) and demographic variables at the refined spatial scale of administrative units ($N=305$) and the coarser scale of counties ($N=62$) for the autumn and winter waves of the 1918–1919 pandemic in England and Wales. The strength of the relationship is given by the Spearman correlation coefficient and the corresponding *p* value. After the Bonferroni correction for multiple comparisons, correlations were deemed significant (indicated in italics) when $p < 0.0025$ (0.05/20).

variable	autumn wave		winter wave	
	units ($N^a=264$)	counties ($N^a=62$)	units ($N^a=209$)	counties ($N^a=54$)
population size	urban	0.09 ($p=0.18$)	0.31 (0.02)	urban	0.18 ($p=0.02$)	0.3 ($p=0.03$)
	rural	0.32 ($p=0.01$)		rural	0.22 ($p=0.13$)	
population density	urban	0.004 (0.96)	0.29 ($p=0.03$)	urban	0.05 ($p=0.52$)	0.22 ($p=0.11$)
	rural	0.21 ($p=0.12$)		rural	0.18 ($p=0.22$)	
average number of people per dwelling	urban	0.09 ($p=0.22$)	0.24 ($p=0.06$)	urban	−0.03 ($p=0.70$)	*0.42 ($p=0.002$)*
	rural	0.13 ($p=0.33$)		rural	0.36 ($p=0.01$)	
average number of people per room	urban	−0.07 ($p=0.30$)	0.22 ($p=0.08$)	urban	0.09 ($p=0.27$)	0.38 ($p=0.005$)
	rural	0.19 ($p=0.14$)		rural	0.36 ($p=0.01$)	

[a] *R* could be estimated in 87% of the 305 administrative units and 100% of the 62 counties during the autumn wave, and in 69% and 87% of the administrative units and counties during the winter wave, respectively.

(95% CI: 1.36–1.43). For the winter wave, we estimated an overall mean *R* of 1.35 (95% CI: 1.33–1.37) in administrative units, with similar values at the county and national levels. Higher *R* estimates were found for a longer serial interval (approx. 1.9 for autumn and approx. 1.7 for winter).

Assuming the extreme case of fixed latent and infectious periods, we can give an upper bound on the estimates of *R* (Wallinga & Lipsitch 2007). In this sensitivity analysis, the *R* estimates increased only marginally, by 0.05 and 0.2 on average, when using the generation intervals of 3 and 6 days, respectively.

Overall, the autumn wave showed higher transmissibility than the winter wave, with 62% of the administrative units experiencing a reduction of transmissibility from autumn to winter (see figure S2 in the electronic supplementary material for maps of *R* estimates). There was no correlation between the reproduction numbers in the autumn and winter waves.

(ii) *Heterogeneity in transmissibility and relationship with socio-demographic factors*
Geographical heterogeneity in influenza transmissibility was statistically significant in the autumn (ANOVA, $p=0.02$) but not in the winter (ANOVA, $p=0.11$). There was no difference in transmissibility between urban and rural areas for either wave (Wilcoxon test, $p > 0.38$).

The associations between *R* estimates and socio-demographic factors were weak to moderate, with the highest correlation estimated at 0.42 ($p=0.002$,

table 3). It should be noted that there was a moderate but significant correlation between the transmissibility in the winter wave and the two measures of residential crowding at the county level (number of people per room and number of people per dwelling). However, no socio-demographic factor was consistently associated with transmissibility at both spatial scales or for both pandemic waves.

(b) *Cumulative death rates in the autumn and winter waves*
(i) *Overall patterns, rural and urban areas*
The average influenza mortality rate was 0.27% (s.d.=0.26%) in the autumn and 0.1% in the winter (s.d.=0.09%), with 94% of the administrative units experiencing decreasing rates from autumn to winter (maps of mortality rates are shown in figure S3 of the electronic supplementary material). Death rates in the autumn and winter waves were weakly correlated ($\rho=0.19$, $p=0.007$). A striking difference was found with urbanization, with 30–40% higher per capita death rates in urban areas than rural areas, for both pandemic waves (Wilcoxon test: autumn wave, $p=0.007$; winter wave, $p=0.0001$; figure S4 in the electronic supplementary material).

(ii) *Relationship with socio-demographic factors*
Table 4 presents the correlations between influenza death rates and demographics for urban and rural units, counties and pandemic waves. The strongest relationship was
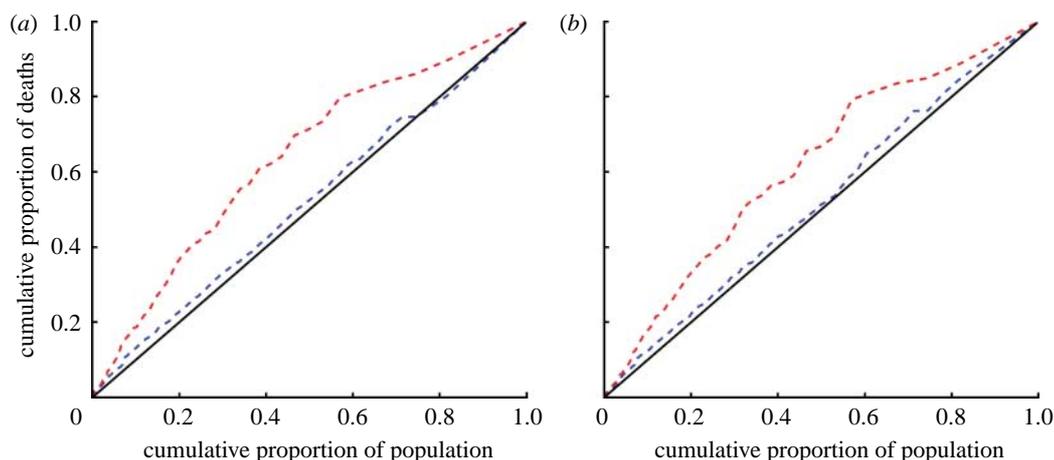
Figure 2. Death rates and population size: the Lorenz curves illustrate the distribution of the total number of influenza deaths as a function of population size for the (*a*) autumn and (*b*) winter waves, at the administrative unit level. Red curves, rural areas; blue curves, urban areas; black lines (first diagonal) illustrate a constant distribution of death rates (no heterogeneity in death rates).

Table 4. Relationship between influenza death rates and demographic variables at the refined spatial scale of administrative units ($N=305$) and the coarser scale of counties ($N=62$) for the autumn and winter waves of the 1918–1919 pandemic in England and Wales. The strength of the relationship is given by the Spearman correlation coefficient and the corresponding p value. After the Bonferroni correction for multiple comparisons, correlations were deemed significant (indicated in italics) when $p<0.0025$ (0.05/20).

variable		autumn wave		winter wave	
		units ($N=305$)	counties ($N=62$)	units ($N=305$)	counties ($N=62$)
population size	urban	-0.13 ($p=0.03$)	0.12 ($p=0.35$)	-0.01 ($p=0.85$)	0.29 ($p=0.02$)
	rural	*-0.46 ($p=0.0003$)*		-0.36 ($p=0.005$)	
population density	urban	0.16 ($p=0.005$)	0.06 ($p=0.62$)	0.08 ($p=0.23$)	0.20 ($p=0.12$)
	rural	-0.21 ($p=0.11$)		-0.13 ($p=0.33$)	
average number of people per dwelling	urban	-0.01 ($p=0.81$)	-0.12 ($p=0.35$)	-0.15 ($p=0.02$)	0.21 ($p=0.10$)
	rural	-0.1 ($p=0.45$)		0.11 ($p=0.42$)	
average number of people per room	urban	0.10 ($p=0.09$)	0.15 ($p=0.26$)	*0.26 ($p<0.0001$)*	*0.41 ($p=0.001$)*
	rural	-0.07 ($p=0.59$)		0.007 ($p=0.96$)	

found with population size in rural settings, where the highest death rates were observed in the areas with the lowest population size—a pattern that was consistent for both pandemic waves. By contrast, there was no association seen for urban areas, and no relationship between death rates and population density. Death rates were weakly associated with an indicator of residential crowding (the average number of people per room) at both spatial scales in cities and towns, but only for the winter wave. Stepwise multivariate regression identified population size and urbanization as the only predictors of death rates in the autumn and winter waves.

(iii) *Modelling death rates as a function of urbanization and population size*
The relationship between the cumulative number of influenza deaths in a geographical area and the population size of the area can be visualized through the Lorenz curves (figure 2). The curve for rural areas was far from the diagonal for both pandemic waves (Gini index between 0.23 and 0.27), thus confirming a systematic relationship between death rates and population size in rural locations. By contrast, there was no association in urban areas (Gini index <0.05). Interestingly, variability in death

rates entirely disappeared when the data were aggregated at the county level (Gini index ~ 0).

We also characterized the relationship between population size and cumulative influenza deaths using a power-law function ($D \propto N^g$), stratified by urban and rural areas (figure 3). For rural areas, the exponent g estimates ranged between 0.71 and 0.77 (significantly below 1.0), whereas these estimates were approximately 1.0 for cities and towns. These estimates suggest that, in rural settings, smaller population units suffered a disproportionately large per capita mortality burden, whereas there was little variation in death rates across cities and towns. In line with the Lorenz curve analysis, heterogeneities disappeared at the scale of counties, and death rates became nearly independent of population size.

(c) *Timing of pandemic onset and socio-demographic factors*
Units with large population size experienced an early pandemic onset, both for the autumn and winter waves (figure 4). There was an approximately three-week difference on average between onset in the high and low population units in the autumn wave, while the winter wave was more synchronized, with an average difference of one week. Population size was well correlated with the
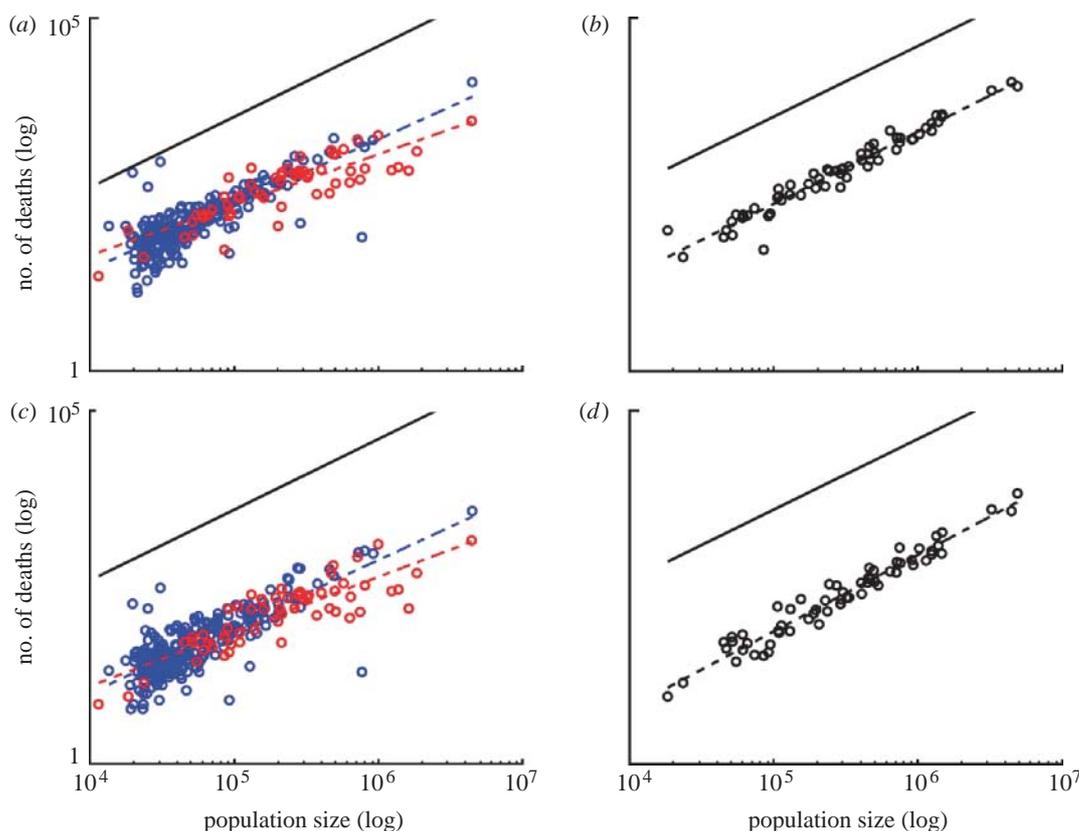
Figure 3. Power-law relationship between the total number of influenza deaths and population size, following $D \sim N^g$ (D, deaths; N, population size; g, exponent), by pandemic wave and spatial scale. The dashed lines represent the best linear fits to the data in log-log scale. The solid black line with a slope of one is shown as a reference to illustrate constant death rates with population size. (*a*) 305 units, autumn; (*b*) 62 counties, autumn; (*c*) 305 units, winter; (*d*) 62 counties, winter. At the refined spatial scale of administrative units, the slope is 'sublinear' for rural areas (red circles, $n = 58$; $g = 0.71$ (95% CI: 0.60–0.83) in autumn and $g = 0.77$ (0.64–0.90) in winter, significantly below 1.0). By contrast for urban areas, the slope is not different from one (blue circles, $n = 247$; $g = 0.92$ (0.84–1.0) in autumn and $g = 0.94$ (0.85–1.03) in winter. At the county level, the power-law exponent was estimated to be 1.02 (0.96–1.08) and 1.08 (1.01–1.14) for the autumn and winter waves, respectively.

timing of the onset of the autumn and winter waves ($-0.55 \le$ Spearman $\rho \le -0.50$, $p < 0.001$). The association was maintained when the data were stratified by rural and urban units, which suggests that population size, rather than urbanization, was the predominant factor driving timing of pandemic onset. Population size was the only socio-demographic factor that independently predicted the timing of the onset for both pandemic waves.

4. DISCUSSION

We explored the geographical and demographic patterns of transmissibility and mortality during the autumn and winter waves of the 1918–1919 influenza pandemic in England and Wales, using weekly epidemiological time series at various levels of spatial aggregation. To the best of our knowledge, this is the first population study of the 1918–1919 influenza pandemic at such a detailed spatial resolution. Our estimates of the reproduction number ranged between 1.3 and 1.9, with a low level of spatial variation across geographical units, which was only marginally associated with socio-demographic factors. By contrast, there was a marked spatial heterogeneity in death rates, which was linked to urbanization, with 30–40% higher rates in cities and towns than in rural areas on average. We evidenced further variations in death rates across rural settings, where smaller population units fared worse. Urban death rates did not vary with population size.

There are several caveats and assumptions in our study. As in several previous studies (Mills *et al.* 2004; Viboud *et al.* 2004, 2006*a,b*), we assumed that influenza mortality was a good proxy for disease incidence and was appropriate for the estimation of transmissibility. Our analyses were conducted with the assumption that reporting and coding of death certificates were homogeneous across England and Wales and over the course of the pandemic (vital statistics had been in place since 1837 in England and Wales, and all deaths were medically certified in 1918). One may expect that more remote locations would have poorer reporting or coding of deaths; however, we found higher death rates in rural units with lower population sizes, which argues against such bias. Finally, as in previous mortality studies (Ministry of Health 1920; Smallman-Raynor *et al.* 2004; Johnson 2006), we relied on the administrative divisions of England and Wales in 1918, which are not necessarily the most meaningful spatial units for disease dynamics.

We identified spatial variations in the R estimates for the autumn wave at the county level; these variations were not simply the result of measurement error, but there was no obvious association with the socio-demographic factors under study. For comparison, a previous analysis of the 1918 flu pandemic in 45 US cities found that the transmissibility of the autumn wave was weakly correlated with population density (Mills *et al.* 2004), and a study of measles in 60 cities in England and Wales during the pre-vaccination era
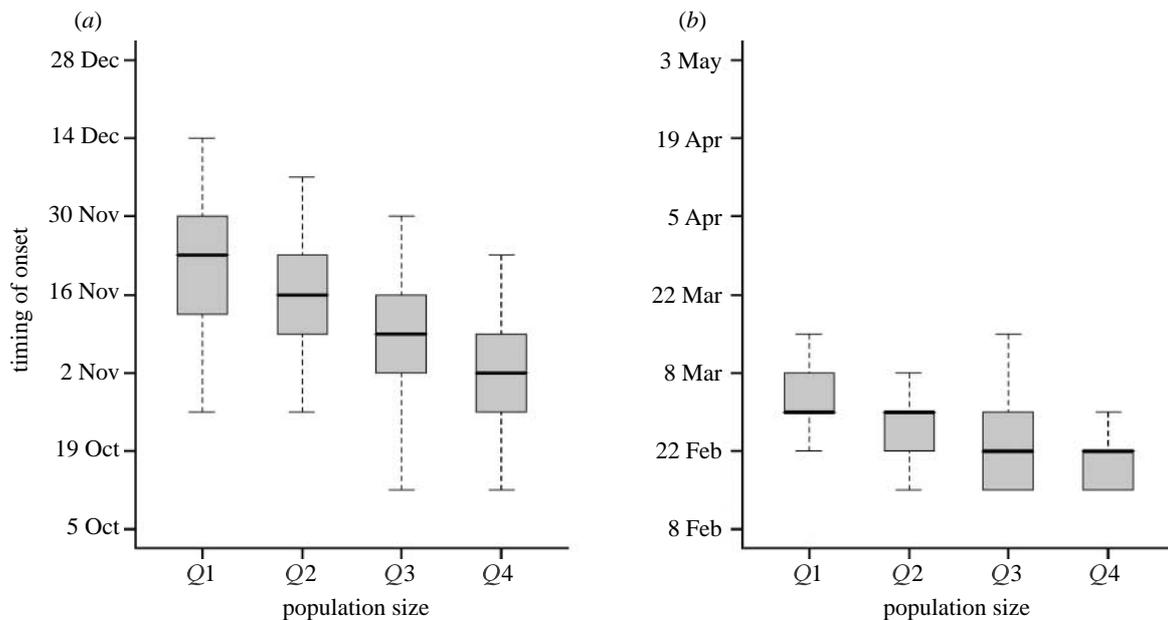
Figure 4. Timing of pandemic onset in the (*a*) autumn (1918) and (*b*) winter (1919) wave in the 305 administrative units of England and Wales, as a function of population size. Q1–Q4 represent the quartiles of population size (Q1, lowest; Q4, highest).

found that transmissibility was not associated with population size (Bjornstad *et al.* 2002). It is interesting that influenza and measles transmissibility appear nearly invariant across a wide range of population sizes and densities, and despite the differences in social connectivity patterns in urban and rural settings. It would be interesting to study this relationship further for seasonal flu epidemics or geographical regions with greater variations in socio-demographic factors.

We found similar transmissibility estimates in the autumn and winter pandemic waves for three different levels of spatial aggregation, including the administrative units, the counties and the entire region of England and Wales. On average, our autumn estimates using a generation interval of 6 days were slightly lower than estimates for the autumn 1918 pandemic wave in the US cities (Mills *et al.* 2004; Chowell *et al.* 2007*a*) and higher than estimates for recent seasonal influenza epidemics (Chowell *et al.* 2007*b*). While the estimates were not substantially changed by the spatial aggregation of the data or by relaxing assumptions on the distribution of the latent and infectious periods, they did change with different assumptions on the duration of the generation interval. Hence, the generation interval is clearly a key parameter with considerable residual uncertainty (Ferguson *et al.* 2006; Andreasen *et al.* in press).

Our results indicate that areas with larger populations experienced early pandemic onset in the autumn and winter waves, thus suggesting a hierarchical spread of influenza driven by large population centres of England and Wales in 1918–1919 (see also Smallman-Raynor *et al.* 2004). This pattern is reminiscent of seasonal influenza epidemics in the USA (Viboud *et al.* 2006*b*). Despite these observed variations in pandemic timing, it is noteworthy that the 1918–1919 pandemic as a whole was more synchronous in the administrative units of England and Wales than in US cities. The relative synchrony of the pandemic in England and Wales is likely to be explained by strong population mixing and the small geographical

extent of this region (Ferguson *et al.* 2006; Bootsma & Ferguson 2007).

Past research has produced conflicting results on the association between the disease burden of the 1918–1919 pandemic and socio-demographic or geographical factors. In a study of four British cities, no association was found between influenza attack rates and residential crowding (Ministry of Health 1920). Furthermore, influenza-related mortality rates were moderately associated with measures of density and baseline pre-pandemic mortality rates in 45 US cities (Pearl 1921; Bootsma & Ferguson 2007). Similarly, in this study, we did not find any obvious association between death rates and measures of population density or residential crowding, or between death rates and pre-pandemic infant mortality rates (correlation approx. 0.30, not shown). Other studies, by contrast, have reported a strong effect of socio-demographic characteristics on the 1918–1919 pandemic mortality rates, including per capita income and indicators of wealth such as apartment size (Mamelund 2006; Murray *et al.* 2006). These conflicting results may stem from the differences in the spatial scale considered (ranging from household to country) and the choice of socio-demographic indicators.

The most important factor associated with death rate in our study was urbanization, with cities and towns experiencing approximately 30–40% higher death rates than rural areas during both pandemic waves. Although the reasons for this pattern remain unclear, a similar mortality pattern has been evidenced in New Zealand (McSweeny *et al.* 2007). Differential exposure to tuberculosis could have played a role, as this disease has been put forward as a predisposing factor for morbidity and mortality during the 1918 influenza pandemic (Pearl 1919; Noymer & Garenne 2000), and tuberculosis was traditionally more prevalent in urban than rural settings. Alternative hypotheses include remoteness and greater social distancing in rural areas (McSweeny *et al.* 2007), although in this case, we would perhaps expect lower transmissibility in rural areas.

Further differences with urbanization were uncovered in our study, in that death rates varied with population size in rural areas, but not in cities and towns. Disparities in access and organization of health care in rural areas with small populations, which fared worse during the pandemic, could explain these geographical patterns (McSweeny *et al.* 2007).

It is interesting that important residual variability in transmissibility estimates and death rates remained unexplained in our study. We think it unlikely that local differences in population age structure or public health interventions could have played a role. Children are believed to drive disease spread locally within cities or small communities (Monto 1999), although they probably do not disseminate infection over large distances (Viboud *et al.* 2006b). In our study, the proportion of the population comprising children did not explain geographical variations in transmissibility or death rates of the 1918–1919 pandemic (not shown). Although the 1918–1919 pandemic virus was characterized by an unusual age pattern of deaths concentrated among young adults (Olson *et al.* 2005), the local differences in the proportion of young adults did not explain variations in death rates in our study (not shown). Public health interventions are irrelevant in the context of England and Wales, as no intervention was implemented in 1918–1919, in contrast to the USA (Bootsma & Ferguson 2007).

The existence of cross-immunity between the viruses circulating in successive waves of the 1918–1919 pandemic has been hypothesized (Anon. 1919; Andreasen *et al.* in press), and may affect the geographical patterns described in this study. The cross-immunity hypothesis relies on anecdotal evidence supporting the idea that individuals who were infected with influenza during the summer escaped clinical illness or experienced only mild respiratory symptoms in the autumn (Anon. 1919). Accordingly, Scandinavia experienced a large influenza morbidity wave in the summer of 1918, followed by a second wave of low mortality impact in the autumn, relative to other countries (Andreasen *et al.* in press). We were not able to fully investigate the impact of differential exposure to influenza during the summer of 1918 in England and Wales, due to the lack of morbidity data and extremely low death counts for this period. However, we found a weak but significant correlation between mortality rates in the summer and autumn waves, which is consistent with the cross-immunity hypothesis (Spearman $\rho = 0.2$, $p < 0.001$). The possibility of cross-immunity also suggests that our rather low R estimates for the autumn and winter pandemic waves do not represent the maximum transmissibility of the 1918 influenza virus in an entirely naive population (the so-called R_0), which may have been as high as 5.4 (Andreasen *et al.* in press).

Overall, this analysis of the 1918–1919 pandemic in England and Wales reveals that high-resolution spatial data at the level of cities, towns and the surrounding rural areas are key to detecting heterogeneity in influenza transmissibility and death rates. Influenza transmissibility results from a complex combination of pre-existing immunity patterns, population mixing and viral strain characteristics, while mortality is also affected by health care and socio-demographic conditions. Our results suggest that population size and urbanization played a role in the geographical patterns of the 1918–1919 pandemic in England and Wales. Whether these patterns can be generalized to contemporary seasonal influenza epidemics and future pandemics is an interesting topic for future research.

REFERENCES

A Vision of Britain through time 2006 Reports of the 1921 census. See http://www.visionofbritain.org.uk/census/index.jsp.

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans*. New York, NY: Oxford University Press.

Andreasen, V., Viboud, C. & Simonsen, L. In press. Epidemiologic characterization of the summer wave of the 1918 influenza pandemic in Copenhagen: implications for pandemic control strategies. *J. Infect Dis.*

Anon. 1919 The influenza pandemic. *Lancet* **1919**, 386–387.

Bjornstad, O., Finkenstadt, B. & Grenfell, B. 2002 Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecol. Monogr.* **72**, 169–184.

Bootsma, M. C. & Ferguson, N. M. 2007 The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proc. Natl Acad. Sci. USA* **104**, 7588–7593. (doi:10.1073/pnas.0611071104)

Cauchemez, S., Carrat, F., Viboud, C., Valleron, A. J. & Boelle, P. Y. 2004 A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Stat. Med.* **23**, 3469–3487. (doi:10.1002/sim.1912)

Chowell, G., Ammon, C. E., Hengartner, N. W. & Hyman, J. M. 2006 Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: assessing the effects of hypothetical interventions. *J. Theor. Biol.* **241**, 193–204. (doi:10.1016/j.jtbi.2005.11.026)

Chowell, G., Nishiura, H. & Bettencourt, L. M. 2007a Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *J. R. Soc. Interface* **4**, 155–166. (doi:10.1098/rsif.2006.0161)

Chowell, G., Miller, M. A. & Viboud, C. 2007b Seasonal influenza in the United States, France, and Australia: transmission and prospects for control. *Epidemiol. Infect.* (doi:10.1017/s0950268807009144)

Diekmann, O. & Heesterbeek, J. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. New York, NY: Wiley.

Ferguson, N. M., Cummings, D. A. T., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S. & Burke, D. S. 2005 Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209–214. (doi:10.1038/nature04017)

Ferguson, N. M., Cummings, D. A., Fraser, C., Cajka, J. C., Cooley, P. C. & Burke, D. S. 2006 Strategies for mitigating an influenza pandemic. *Nature* **442**, 448–452. (doi:10.1038/nature04795)

Gani, R., Hughes, H., Fleming, D. M., Griffin, T., Medlock, J. & Leach, S. 2005 Potential impact of antiviral drug use during influenza pandemic. *Emerg. Infect. Dis.* **11**, 1355–1362.

Green, C. G., Krause, D. & Wylie, J. 2006 Spatial analysis of *Campylobacter* infection in the Canadian province of Manitoba. *Int. J. Health Geogr.* **5**, 2. (doi:10.1186/1476-072X-5-2)

Hatchett, R. J., Mecher, C. E. & Lipsitch, M. 2007 Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proc. Natl Acad. Sci. USA* **104**, 7582–7587. (doi:10.1073/pnas.0610941104)

Johnson, N. 2006 *Britain and the 1918–19 influenza pandemic: a dark epilogue.* London, UK; New York, NY: Routledge.

Kerani, R. P., Handcock, M. S., Handsfield, H. H. & Holmes, K. K. 2005 Comparative geographic concentrations of 4 sexually transmitted infections. *Am. J. Public Health* **95**, 324–330. (doi:10.2105/AJPH.2003.029413)

Lee, W. C. 1997 Characterizing exposure-disease association in human populations using the Lorenz curve and Gini index. *Stat. Med.* **16**, 729–739. (doi:10.1002/(SICI)1097-0258(19970415)16:7<729::AID-SIM491>3.0.CO;2-A)

Lipsitch, M. *et al.* 2003 Transmission dynamics and control of severe acute respiratory syndrome. *Science* **300**, 1966–1970. (doi:10.1126/science.1086616)

Longini Jr, I. M., Halloran, M. E., Nizam, A. & Yang, Y. 2004 Containing pandemic influenza with antiviral agents. *Am. J. Epidemiol.* **159**, 623–633. (doi:10.1093/aje/kwh092)

Low, R. B. 1920 Report on the pandemic of influenza 1918–19: Europe and Western Hemisphere. Reports on public health and medical subjects no. 4, Her Majesty's Sationery Office, London.

Mamelund, S. E. 2006 A socially neutral disease? Individual social class, household wealth and mortality from Spanish influenza in two socially contrasting parishes in Kristiania 1918–19. *Soc. Sci. Med.* **62**, 923–940. (doi:10.1016/j.socscimed.2005.06.051)

Massad, E., Burattini, M. N., Coutinho, F. A. & Lopez, L. F. 2007 The 1918 influenza A epidemic in the city of Sao Paulo, Brazil. *Med. Hypotheses* **68**, 442–445. (doi:10.1016/j.mehy.2006.07.041)

McSweeny, K., Colman, A., Fancourt, N., Parnell, M., Stantiall, S., Rice, G., Baker, M. & Wilson, N. 2007 Was rurality protective in the 1918 influenza pandemic in New Zealand? *N. Z. Med. J.* **120**, U2579.

Mills, C. E., Robins, J. M. & Lipsitch, M. 2004 Transmissibility of 1918 pandemic influenza. *Nature* **432**, 904–906. (doi:10.1038/nature03063)

Ministry of Health 1920 Supplement to the Eighty-First Annual Report of the Registrar-General-Report on the mortality from influenza in England and Wales during the epidemic of 1918–19. Reports on public health and medical subjects, London, pp. 48–80.

Monto, A. S. 1999 Interrupting the transmission of respiratory tract infections: theory and practice. *Clin. Infect. Dis.* **28**, 200–204.

Murray, C. J., Lopez, A. D., Chin, B., Feehan, D. & Hill, K. H. 2006 Estimation of potential global pandemic influenza mortality on the basis of vital registry data from the 1918–20 pandemic: a quantitative analysis. *Lancet* **368**, 2211–2218. (doi:10.1016/S0140-6736(06)69895-4)

Neter, J. & Wasserman, W. 1974 *Applied linear statistical models.* Boston, MA: Richard D. Irwin, Inc.

Nishiura, H. 2007 Time variations in the transmissibility of pandemic influenza in Prussia, Germany, from 1918–19. *Theor. Biol. Med. Model* **4**, 20. (doi:10.1186/1742-4682-4-20)

Noymer, A. & Garenne, M. 2000 The 1918 influenza epidemic's effects on sex differentials in mortality in the United States. *Popul. Dev. Rev.* **26**, 565–581. (doi:10.1111/j.1728-4457.2000.00565.x)

Olson, D. R., Simonsen, L., Edelson, P. J. & Morse, S. S. 2005 Epidemiological evidence of an early wave of the 1918 influenza pandemic in New York City. *Proc. Natl Acad. Sci. USA* **102**, 11 059–11 063. (doi:10.1073/pnas.0408290102)

Pearl, R. 1919 Preliminary notes on the incidence of epidemic influenza among the actively tuberculous. *Q. Publ. Am. Stat. Assoc.* **16**, 536–540.

Pearl, R. 1921 Influenza studies: further data on the correlation of explosiveness of outbreak of the 1918 epidemic. *Public Health Rep.* **36**, 273–298.

Sertsou, G., Wilson, N., Baker, M., Nelson, P. & Roberts, M. G. 2006 Key transmission parameters of an institutional outbreak during the 1918 influenza pandemic estimated by mathematical modelling. *Theor. Biol. Med. Model* **3**, 38. (doi:10.1186/1742-4682-3-38)

Smallman-Raynor, M., Johnson, N. & Cliff, A. D. 2004 The spatial anatomy of an epidemic: influenza in London and the county boroughs of England and Wales, 1918–1919. *Trans. Inst. Br. Geogr.* **27**, 452–470. (doi:10.1111/1475-5661.00065)

Viboud, C., Boelle, P. Y., Pakdaman, K., Carrat, F., Valleron, A. J. & Flahault, A. 2004 Influenza epidemics in the United States, France, and Australia, 1972–1997. *Emerg. Infect Dis.* **10**, 32–39.

Viboud, C., Tam, T., Fleming, D., Handel, A., Miller, M. A. & Simonsen, L. 2006*a* Transmissibility and mortality impact of epidemic and pandemic influenza, with emphasis on the unusually deadly 1951 epidemic. *Vaccine* **24**, 6701–6707. (doi:10.1016/j.vaccine.2006.05.067)

Viboud, C., Bjornstad, O. N., Smith, D. L., Simonsen, L., Miller, M. A. & Grenfell, B. T. 2006*b* Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451. (doi:10.1126/science.1125237)

Wallinga, J. & Lipsitch, M. 2007 How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B* **274**, 599–604. (doi:10.1098/rspb.2006.3754)

Woolhouse, M. E. *et al.* 1997 Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc. Natl Acad. Sci. USA* **94**, 338–342. (doi:10.1073/pnas.94.1.338)