THE ROYAL SOCIETY

The geographical spread of influenza

Eric Bonabeau[1], **Laurent Toubiana**[2] **and Antoine Flahault**[2]

[1]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*
[2]*INSERM-U444, B3E-ISARS, CHU Saint-Antoine, 27 rue Chaligny, 75012 Paris, France*

How infectious diseases spread in space within one cycle of an epidemic is an important question that has received considerable theoretical attention. There are, however, few empirical studies to support theoretical approaches, because data are scarce. Weekly reports obtained since 1984 from a network of general practitioners spanning the entire French territory allows the analysis of the spatio-temporal dynamics of influenza over a fine spatial scale. This analysis indicates that diffusion over long distances, possibly due to global transportation systems, is so quick that homogeneous global mixing occurs before the epidemic builds up within infected patches. A simple model in which the total number of cases is given by the empirical time-series and cases are randomly assigned to patches according to the population weight of the patches exhibits the same spatio-temporal properties as real epidemic cycles: homogeneous mixing models constitute appropriate descriptions, except in the vicinity of the epidemic's peak, where geographic heterogeneities play a role.

Keywords: epidemics; influenza; geographical spread; global mixing

1. INTRODUCTION

Models developed by epidemiologists are aimed at predicting epidemics or evaluating the efficiency of vaccination (Kermack & McKendrick 1927; May & Anderson 1984; Anderson & May 1991; Anderson 1994; Hethcote 1978; Hethcote & Van Ark 1986; Mollison 1977; Grenfell *et al*. 1995; Post *et al*. 1983). A number of models incorporate real data, some among these include heterogeneities in the form of age structure or interaction networks (Anderson 1994; Hethcote 1978; Post *et al*. 1983; Hethcote & Van Ark 1986; Mollison 1977; Grenfell *et al*. 1995; Woolhouse *et al*. 1997; Bolker & Grenfell 1993, 1996; Ferguson *et al*. 1996), but few deal with the spatial component of epidemics (Baroyan *et al*. 1971; Cliff 1995; Cliff *et al*. 1986; Murray & Cliff 1975; Noble 1974; Cliff & Haggett 1988) despite the importance of understanding the geographical spread of infectious diseases. This can be understood as it is hard to find long and reliable epidemic time-series, and virtually impossible to find good spatio-temporal data with fine-grained sampling scales. A remarkable data set in that respect are the UK measles incidence case records, which span 40 years at more than 2000 locations, with up to several million cases reported per year. This data set has been extensively studied and remarkable results have been found regarding the often subtle but important effects of space (Murray & Cliff 1975; Grenfell *et al*. 1995; Bolker & Grenfell 1996).

The existence of a network of general practitioners (GPs) in France (Valleron *et al*. 1986), evenly distributed over the French territory since 1984, allowed us to obtain good-quality data (detailed reports exist for detected cases) on a weekly basis on a relatively small spatial scale (20 km). The number of GPs involved in the sentinel network has increased significantly since the network was launched, now representing 1% of all French GPs. Consequently, the size of the influenza data set, which is the focus of the present study, reaches several thousand reports within an epidemic cycle of about 15 weeks. Results presented in this paper are given for all epidemic cycles between 1987 and 1995. One question of particular interest is whether or not geographical space, including heterogeneities in population distribution, is relevant to the spread of the epidemic within a cycle. If space is not relevant, global mixing models give accurate descriptions of the epidemic process. If, on the other hand, space does play a role, spatial models incorporating local mixing and locally density-dependent dynamics, may be more appropriate. In terms of immunization, the two descriptions have different consequences (May & Anderson 1984; Anderson & May 1991; Bolker & Grenfell 1996) (uniform versus differential immunization). There has been some evidence in the past that space is indeed relevant to the spread of epidemics, and that epidemics spread with spatial waves from an initial epidemic centre (Cliff 1995; Cliff *et al*. 1986; Murray & Cliff 1975; Noble 1974; Cliff & Haggett 1988). There is also some evidence that the local dynamics of an epidemic, characterized for example by its force of infection or basic reproduction number, depend on heterogeneities in population density (May & Anderson 1984; Anderson & May 1991), although the topic remains somewhat controversial and the dependence may be complex. Transportation systems have changed in the last 30 years, not only promoting exchanges among countries, but also within countries. Although global exchanges did exist 30 years ago (for example, retrospective modelling of the 1968 Hong Kong pandemic suggested that the virus had diffused rapidly, taking less than two years, through a worldwide network of interconnected cities (Rvachev & Longini 1985)), they may have been seriously amplified to the point where the time-scale of global mixing is so short that it dominates

Received 9 July 1998 *Accepted* 20 July 1998

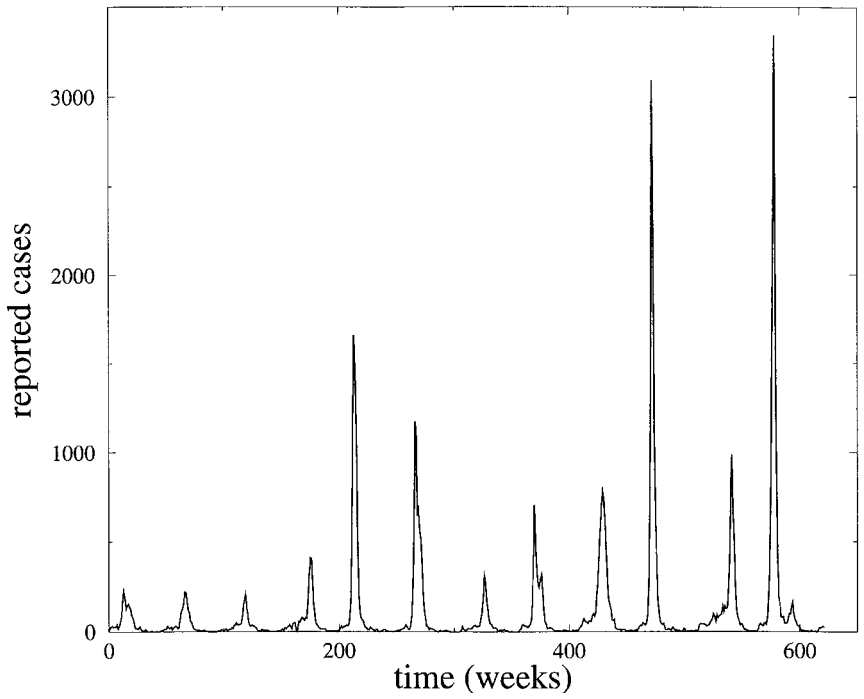

Figure 1. Time-series of the number of cases of influenza reported by the French network of GPs (Sentinelles) between 1984 and 1995. Twelve peaks can be observed. Only the last eight cycles are used in the study. Definition of a case: sudden appearance of fever $>39\,^{\circ}$C, myalgia and respiratory problems. A total of 166 infecteds have been tested for an influenza virus during the 1994–1995 epidemic: 94% had type A viruses (53% H3N2 and 41% H1N1), and 6% had a type B virus.

over local dynamics. This also points to the crucial problem of the speed at which a vaccine should be developed in response to a new pandemic.

2. CORRELATION ANALYSIS

Figure 1 shows the weekly time-series of the total number of cases reported by the GP network over the entire territory between 1984 and 1995. The first four cycles that can be seen in figure 1 will not be used in our study because of the imperfect coverage of the French territory by the network at that time. In order to study the influence of space on the spread of the epidemic within each period of the oscillatory incidence pattern, spatial correlations have been computed. Spatial data are binned in square patches of linear size s km, hereafter called s-patches. Let $N_s(r,t)$ be the number of cases reported during week t in the s-patch, centred on location r. Let $n_s(r,t)=N_s(r,t)/P_s(r)$ denote the 'normalized' version of $N_s(r,t)$, where $P_s(r)$ is the total population in the s-patch centred around r: $n_s(r,t)$ represents the per capita incidence data. Figure 2 shows the equal-time density–density autocorrelation function $C_s(h,t)$ of $n_s(r,t)$ at the peaks of the last eight epidemic cycles observed in figure 1. $C_s(h,t)$ is defined by

$$C_s(h,t) = \frac{\langle\langle[n_s(rt)-\bar{n}][n_s(r+y,t)-\bar{n}]\rangle_r\rangle_{|y|=h}}{\sigma^2(n)}$$

where t is the time in weeks, $\langle\cdots\rangle_r$ denotes averaging over r and $\langle\cdots\rangle_{|y|=h}$ averaging over $|y|=h$, $\bar{n}=\langle n_s(r,t)\rangle_r$, is the average spatial value of $n_s(r,t)$ and $\sigma^2(n)=\langle[n_s(r,t)-\bar{n}]^2\rangle_r$ is the spatial variance of $n_s(r,t)$. The spatial distribution of the French population has been obtained from the French Institut Géographique National (last 1990 census survey). The correlations in $n_s(r,t)$ can be reasonably well characterized by $C_s(h,t)<h^{\alpha}$ with $\alpha=0.04\pm0.03$, that is, a flat correlation function. This observation is confirmed when the same quantity, $C_s(h,t)$, is plotted for the whole data set (figure 3). The best global fit of the type for the cloud of points represented in figure 3 is for $\alpha=-0.05\pm0.06$. This result suggests that the distribution of cases is randomly distributed over the territory, with the number of cases in each patch being roughly proportional to the population of the patch. When the correlations of the raw incidence data $N_s(r,t)$, rather than those of its normalized value $n_s(r,t)$, are studied, long-range correlations are obtained (Bonabeau *et al.* 1998): such long-range correlations reflect the spatial structure of the underlying population distribution and not intrinsic spatial properties of the epidemic. It can therefore be misleading to use absolute incidence: per capita incidence data should always be preferred.

3. NULL HYPOTHESIS

A test of the suggestion made in the previous section—that the spread of the epidemic is statistically uniform in geographic space—consists of studying the empirical relationship between the epidemic's spatial extent (number of infected patches) and its incidence (number of cases). The number of new reported cases per patch, $N_s(r,t)$, implicitly includes a threshold procedure: only in those regions where the actual number of cases exceeds some threshold can we expect reports to be issued. The spatial extent provided by the GP network may therefore not be the true extent of the epidemic but rather the number of regions in which the number of cases has exceeded the threshold. On the basis of this information, the following null hypothesis can be tested. Let I_t be the total number of detected cases at time t over the whole territory and P the total population. The null hypothesis assumes that the I_t detected cases are randomly distributed in space among all the regions of linear size s according to their demographic weights—in other words, incidence is merely correlated with local population density.

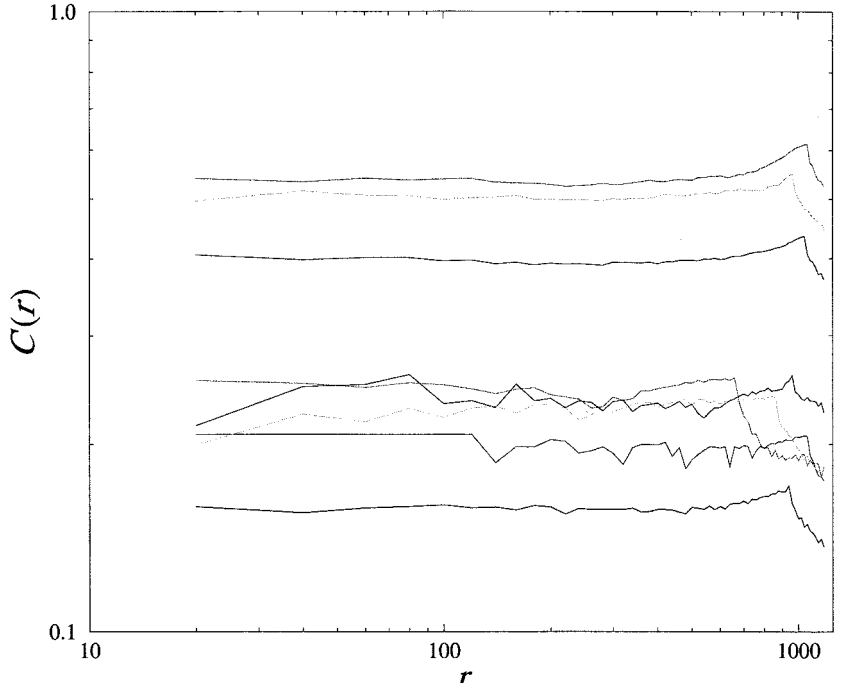

Figure 2. Log–log plot of the spatial equal-time density–density auto-correlation function $C_s(h,t)$ of $n_s(r,t)$, the number of reported cases normalized by the size of the population, at scale $s=20$ km. The eight represented curves correspond to the peaks of the last eight epidemic cycles of figure 1. $C_s(h,t)$ can be reasonably well characterized by $C_s(h,t) \propto h^{\alpha}$ with $\alpha = 0.04 \pm 0.03$.

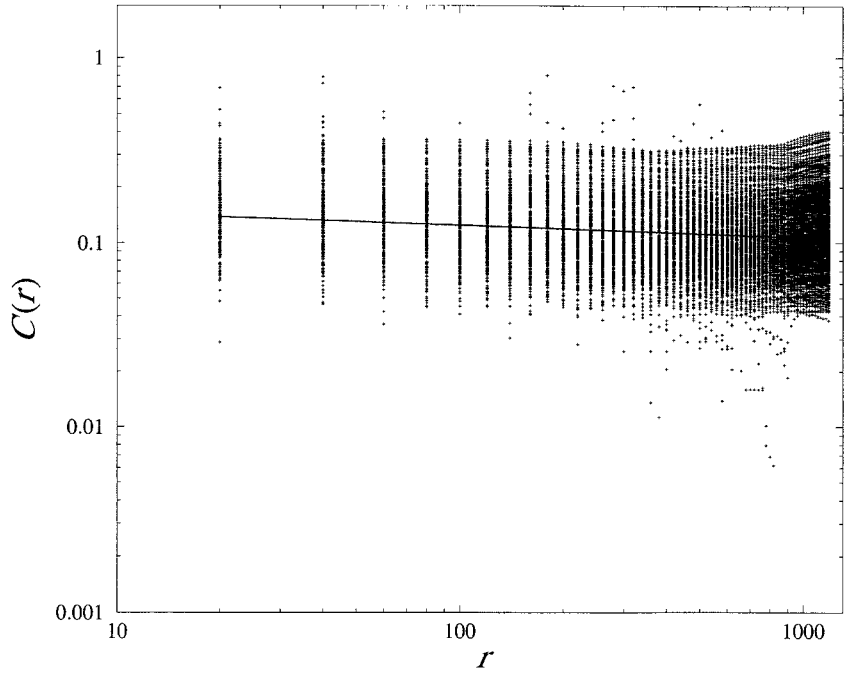

Figure 3. Log–log plot of $C_s(h,t)$ for the whole 1987–1995 data set. The best global fit of the type $C_s(h,t) \propto h^{\alpha}$ for the cloud of points is for $\alpha = -0.05 \pm 0.06$.

In order to simulate this hypothesis, each of the I_t cases is assigned to a zone: a zone i (of linear size s around point ri) is chosen at random, and the case is assigned to zone i with probability $P_s(r)/P$; if zone i is rejected, another zone is picked at random and tested, and this procedure is repeated until all cases are assigned to a zone. The relationship found between the size of the epidemic and its spatial extent observed in the data can be compared to the one obtained with the shuffled data. This method is similar to the surrogate data method used in time-series analysis (Theiler *et al*. 1992): spatially distributing the cases among zones according to their demographic weights leads by construction to the same spatial correlations as those observed in the data.

Figure 4 shows the number of infected patches as a function of the number of reported cases in the data and in 1000 simulations of the null hypothesis. One cannot easily discriminate between the null hypothesis and the data, except at the peak of the epidemic, indicating that the epidemic 'spreads' more rapidly than it 'builds up' within a given patch until it reaches its peak, whereas the real spatial extent of the epidemic saturates at a slightly lower value than the simulated one. In both the data and the simulation, three regimes are observed: (i) before and after the epidemic (number of reported cases <20), where there is a close-to-linear relationship between the spatial extent and the number of cases, (ii) an intermediate phase ($20 <$ number of reported cases < 200) where the number of infected patches increases sublinearly with the number of reported cases, and (iii) the vicinity of the epidemic's peak ($\mathcal{N} > 200$), where the number of infected patches saturates. The small difference between the real and simulated spatial

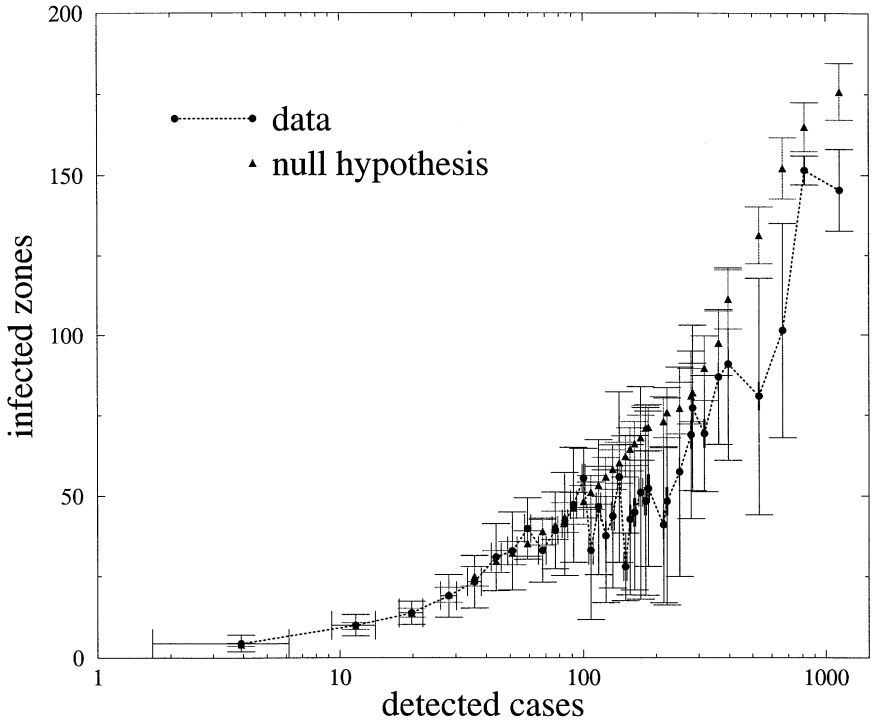

Figure 4. Number of infected patches as a function of the number of reported cases in the data and in 1000 simulations of the null hypothesis. The null hypothesis consists of assigning cases to patches in proportion to their demographic weights. In the empirical data set, error bars along the x-axis correspond to putting I_t into bins and averaging over all data points within each bin, and error bars along the y-axis correspond to averaging over all spatial extents obtained for each bin. The empirical data set aggregates all data between 1987 and 1995. Error bars in the simulated data correspond to averaging over 1000 simulations. The spatial scale for both the data and the simulations is $s=20$ km.

extent in regime (iii) is likely to result from the following factor: in the simulation, the spatial extent saturates as the epidemic covers most of the territory, whereas in reality, it saturates because the epidemic reaches highly populated regions where contact rates may be higher, so that the number of cases increases in these regions more than in other regions.

4. BASIC REPRODUCTION NUMBER

On the basis of the observations in the previous sections, it is tempting to conclude that the data reflect a random distribution of cases according to demographic weight, and that the epidemic propagates mainly through a globally mixing process, possibly because of global transportation systems (Flahault *et al*. 1988, 1994). The dynamics of the global incidence (that is, the total number of cases over the whole territory) are a mean-field dynamics defined mainly at the global level under 'homogeneous mixing'. Density dependence seems to play a role only in the vicinity of the epidemic's peak. When a region becomes infected, the number of cases in that region increases, but the macroscopic spread of the epidemic is more rapid than its microscopic dynamics, so that a mean-field regime is reached before local dynamical heterogeneities can dominate. Only when the epidemic reaches its peak do such heterogeneities come into play, because of the build-up of the epidemic within densely populated infected patches.

If this theory is valid, the basic reproduction number, denoted by R_0, should not vary significantly between patches up to some large population density value, above which R_0 becomes greater. In order to test this prediction, R_0 can be approximated within a discrete time formalism (where 1 unit=1 week) by $R_0 \approx n^{1/T_n}$, where T_n is the number of weeks it takes for the per capita incidence to be multiplied by n (here, $n=2,3$) (Dietz 1993). This approximation results from the fact that the initial steps of the epidemic can be described by a geometric process with parameter R_0. Regions in which the per capita incidence does not double or triple within the whole epidemic cycle are not taken into account. More precisely, the following procedure is applied to estimate R_0.

(i) Each cycle of the oscillatory incidence pattern is studied separately, starting 20 weeks before the peak of the cycle.

(ii) The first time, t_1, when the per capita incidence n_1 becomes greater than or equal to 5×10^{-6} (that is one case for 2×10^5 inhabitants), is reported in a given s-patch ($s=20$ km) is monitored. Both t_1 and n_1 are recorded. The doubling time T_2 is defined by $T_2=t_2-t_1$, where t_2 is the first time, if applicable, when the per capita incidence exceeds $2n_1$. If the per capita incidence does not double within the epidemic cycle, the data corresponding to the patch are discarded.

(iii) The tripling time T_3 is computed in very much the same way as the doubling time T_2. T_3 is defined by $T_3=t_3-t_1$, where t_3 is the first time, if applicable, when the per capita incidence exceeds $3n_1$. If the per capita incidence does not triple within the epidemic cycle, the data corresponding to the patch are discarded.

We found no simple relationship between population density (PD, where PD is the average geographic density within a patch and does not exceed 2500 inhabitants km^{-2}) and R_0, as already suggested by numerous authors (e.g. de Jong *et al*. 1995). However, larger values of R_0 are obtained for larger values of PD. For $0<\mathrm{PD}\leqslant1300$ inhabitants km^{-2}, a cloud of R_0 values ranging from 1.09–1.732 can be observed. For $1300<\mathrm{PD}<2500$ inhabitants km^{-2}, two clearly separated clouds of data points can be observed, which correspond to (i) regions where R_0 is comparable to the one found for $0<\mathrm{PD}\leqslant1300$ inhabitants km^{-2}, and (ii) regions where R_0

is larger ($R_0=2$ or 3). Some high-density patches yield larger R_0 estimates than others: because PD is the average PD within patches, two patches with the same PD may have quite different spatial organizations and contact rates; how the population is spatially distributed within patches can influence R_0. Despite the biases inherent in the computation of R_0, the observed difference in R_0 between some highly populated regions and other regions seems to confirm that the difference between the null hypothesis (§ 3) and the data results from a heterogeneity in contact rates that becomes apparent only when the epidemic reaches densely populated regions.

5. DISCUSSION

In order to be completely sure that the apparent global mixing does not result from the annual flare-up of an already locally circulating infection (influenza can be maintained in small populations), one would need to know the antigenic types of all recorded cases, all the more as influenza is subject to antigenic drift. However, about 15% of all reported cases were tested during the 1994–1995 epidemic cycle and indicate a strong proportion of two type A strains (53% H3N2 and 41% H1N1). This suggests that most of our analysis should be correct: models assuming global homogeneous (and instantaneous, within a week) geographic mixing are appropriate to describe the initial spread of the epidemic; geographic heterogeneities and density dependence play a role only in the few weeks around the epidemic's peak

We thank the sentinel general practitioners who collected the data. The Sentinelles network is part of the French Communicable Diseases Network (FCDN), developed at INSERM U444 in collaboration with the Réseau National de Santé Publique (Public Health Network) and the Direction Générale de la Santé (Ministry of Health). E.B. is supported by the Interval Research fellowship at the Santa Fe Institute.

REFERENCES

Anderson, R. M. 1994 The Croonian Lecture 1994. Populations, infectious disease and immunity: a very nonlinear world. *Phil. Trans. R. Soc. Lond.* B **346**, 457–505.

Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans. Dynamics and control*. Oxford University Press.

Baroyan, O. V., Rvachev, L. A., Basilevsky, U. V., Ezmakov, V. V., Frank, K. D., Rvachev, M. A. & Shaskov, V. A. 1971 Computer modelling of influenza epidemics for the whole country (USSR). *Adv. App. Prob.* **3**, 224–226.

Bolker, B. M. & Grenfell, B. T. 1993 Chaos and biological complexity in measles dynamics. *Proc. R. Soc. Lond.* B **251**, 75–81.

Bolker, B. M. & Grenfell, B. T. 1996 Impact of vaccination on the spatial correlation and persistence of measles dynamics. *Proc. Natn. Acad. Sci. USA* **93**, 12 648–12 653.

Bonabeau, E., Toubiana, L. & Flahault, A. 1998 Evidence for global mixing in real influenza epidemics. *J. Phys.* A **31**, L361–L363.

Cliff, A. D. 1995 Incorporating spatial components into models of epidemic models. In *Epidemic models—their structure and relation to data* (ed. D. Mollison), pp. 119–149. Cambridge University Press.

Cliff, A. D. & Haggett, P. 1988 *Atlas of disease distributions: analytic approaches to epidemiological data*. Oxford, UK: Blackwell.

Cliff, A. D., Haggett, P. & Ord, J. K. 1986 *Spatial aspects of influenza epidemics*. London: Pion.

de Jong, M. C. M., Diekman, O. & Heesterbeek, H. 1995 How does transmission of infection depend on population size? In *Epidemic models—their structure and relation to data* (ed. D. Mollison), pp. 84–94. Cambridge University Press.

Dietz, K. 1993 The estimation of the basic reproduction number for infectious diseases. *Statist. Meth. Med. Res.* **2**, 23–41.

Ferguson, N. M., Nokes, D. J. & Anderson, R. M. 1996 Dynamical complexity in age-structured models of the transmission of the measles virus: epidemiological implications at high levels of vaccine uptake. *Math. Biosci.* **138**, 101–130.

Flahaut, A., Letrait, S., Blin, P., Hazout, S., Ménarerès, J. & Valleron, A.-J. 1988 Modelling the 1985 influenza epidemic in France. *Statist. Med.* **7**, 1147–1155.

Flahault, A., Deguen, S. & Valleron, A.-J. 1994 A mathematical model for the European spread of influenza. *Eur. J. Epidemiol.* **10**, 471–474.

Grenfell, B. T., Kleczkowski, A., Gilligan, C. A. & Bolker, B. M. 1995 Spatial heterogeneity, nonlinear dynamics and chaos in infectious diseases. *Statist. Meth. Med. Res.* **4**, 160–183.

Hethcote, H. W. 1978 An immunization model for a heterogeneous population. *Theor. Popul. Biol.* **14**, 338–349.

Hethcote, H. W. & Van Ark, J. W. 1986 Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation and immunization programs. *Math. Biosci.* **84**, 84–118.

Kermack, W. O. & McKendrick, A. G. 1927 A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* A **115**, 700–721.

May, R. M. & Anderson, R. M. 1984 Spatial heterogeneity and the design of immunization programs. *Math. Biosci.* **72**, 83–111.

Mollison, D. 1977 Spatial contact models for ecological and epidemic spread. *J. R. Statist. Soc.* B **39**, 283–326.

Murray, G. D. & Cliff, A. D. 1975 A stochastic model for measles epidemics in a multi-region setting. *Instit. Br. Geogr.* **2**, 158–174.

Noble, J. V. 1974 Geographic and temporal development of plagues. *Nature* **250**, 726–729.

Post, W. M., DeAngelis, D. L. & Travis, C. C. 1983 Endemic disease in environments with spatially heterogeneous host populations. *Math. Biosci.* **63**, 289–302.

Rvachev, L. A. & Longini, I. M. 1985 A mathematical model for the global spread of influenza. *Math. Biosci.* **75**, 1–22.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Farmer, J. D. 1992 Testing for nonlinearity in time series: the method of surrogate data. *Physica* D **58**, 77–94.

Valleron, A. J., Bouvet, E., Garnerin, P., Ménarerès, J., Heard, I., Letrait, S. & Lefaucheux, J. 1986 A computer network for the surveillance of communicable diseases: the French experiment. *Am. J. Public Hlth* **76**, 1289–1292.

Woolhouse, M. E. J. (and 12 others) 1997 Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc. Natn. Acad. Sci. USA* **94**, 338–342.