

SCIENTIFIC REPORTS



OPEN

Positively interacting strains that co-circulate within a network structured population induce cycling epidemics of *Mycoplasma pneumoniae*

Xu-Sheng Zhang^{1,2}, Hongxin Zhao¹, Emilia Vynnycky^{1,3} & Vicki Chalker¹

Mycoplasma pneumoniae (MP) is considered a common cause of pneumonia, causing about 15–20% of adult community-acquired pneumonia (CAP) and up to 40% of cases in children. It has often been observed that MP epidemics last approximately 1–2 years and occur every 3–7 years, with the dominant strains alternating between epidemics. However, the underlying mechanism by which these cycles and changes in the dominant strains occur remains unclear. The traditional models for the periodicity of MP epidemics neglected two phenomena: structured contact patterns among people and co-circulating strains of MP. We also believe that the two distinctive aspects of MP epidemics: prevalent serotype shifts among epidemics and incidence cycling of MP, are interconnected. We propose a network transmission model that assumes two strains of MP are transmitted within a network structured population and they can interact as secondary infections with primary infections. Our studies show that multiple strains that co-circulate within a network structured population and interact positively generate the observed patterns of recurrent epidemics of MP. Hence our study provides a possible mechanism for the cycling epidemics of MP, and could provide useful information for future vaccine design and vaccine evaluation/monitoring processes.

Mycoplasma pneumoniae (MP) is an “atypical” bacterium that causes acute respiratory infection in humans of all ages. MP is considered a common cause of pneumonia: MP causes about 15–20% adult community-acquired pneumonia (CAP) and up to 40% cases in children; however, not every infected patient actually develops pneumonia^{1–3}. MP infection generally tends to occur more frequently during the summer and autumn months when other respiratory pathogens are less prevalent; but the disease incidence does not appear to be related to season or geography^{4,5}. For example, we also notice that MP infections have been observed to occur more frequently in winter months in England and Wales^{6,7}. Epidemics of MP tend to occur every 3–7 years in the general population^{6–8}. Analysis of laboratory reports of MP infections in England and Wales from 1975 to 2009⁹ has indicated that these epidemics last on average 18 months occurring at approximately four yearly intervals. MP is a polymorphic pathogen^{10,11}; for example, Chalker *et al.*⁶ identified eleven strain types circulating in England and Wales during October 2010 to January 2011. MP strains can be differentiated based on differences in the P1 adhesin gene or in the number of repetitive sequences at a given genomic locus using multilocus variable number tandem repeat analysis (MLVA)^{12,13}. Kenri *et al.*¹⁴ noticed that more than one serotype of MP were circulating within Japanese populations. Kogoj *et al.*¹⁵ observed a shift in the dominant MP strain between two epidemics that occurred in Slovenia in 2006 and 2016. Multiple strains of MP and their co-circulation were also observed in other countries (e.g.^{7,16–19}). Although there are many different isolates and strains, analysis of repetitive elements distributed in

¹Centre for Infectious Disease Surveillance and Control, Public Health England, London, UK. ²Medical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College School of Public Health, London, UK. ³TB Modelling Group, TB Centre, Centre for Mathematical Modelling of Infectious Diseases and Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK. Correspondence and requests for materials should be addressed to X.-S.Z. (email: xu-sheng.zhang@phe.gov.uk)

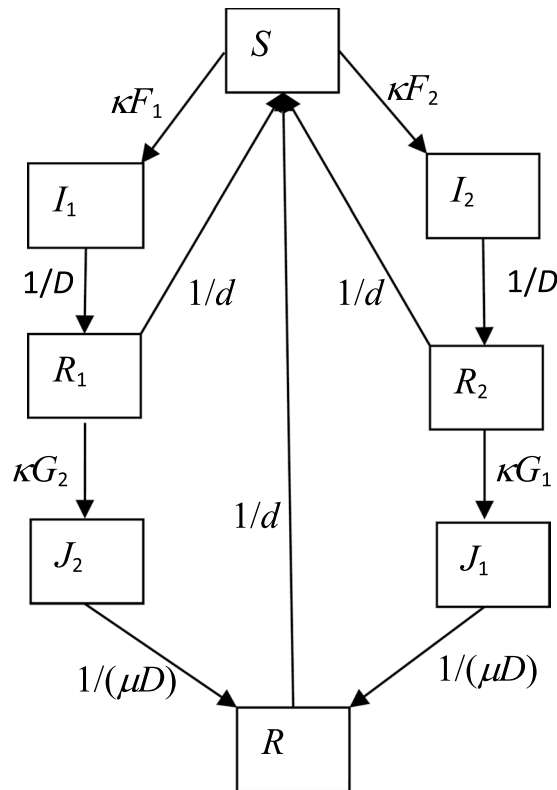


Figure 1. Flow chart of the two-strain SIRS epidemic model. Arrows indicate transitions. Expressions next to arrows show the *per capita* flow rate between compartments. Births and deaths are not shown. Parameter κ is the degree of contacts each person possesses and μ is the effect of primary infection strain on the duration of infection by the secondary strain. Variables F_1 (F_2) and G_1 (G_2) are forces of infection of strain 1 (strain 2) that are defined in equation (4).

variable size and sequence over the genome of MP strains suggested two main types: P1 type 1 and P1 type 2^{14,16,20,21}.

Humans are the sole reservoir of MP and transmission requires close contact. Outbreaks typically occur within closed populations, such as in schools, military premises and prisons. Airborne spread of aerosols and, potentially, indirect contact with contaminated items, may contribute to transmission. The transmissibility of an infectious agent can be estimated by calculating the basic reproduction number (R_0), which is defined as the mean number of secondary infectious cases generated by one primary infectious case introduced into a totally susceptible population²². Using seroprevalence data from a western population, Nguipdop-Djomo *et al.*⁹ estimated R_0 of MP to be 1.7 (95% confidence interval (CI) 1.6–1.9), indicating low transmissibility. The incubation period of MP averages 2 to 3 weeks. The duration of infectiousness is unclear and is commonly estimated to be up to 3 weeks from onset of illness²³. Immunity occurs post infection, but later re-infection with different subtypes is recognized, suggesting the immunity is not lifelong and no strong cross protection between different subtypes^{21,24,25}. The duration of immunity ranges from 2 to 10 years²⁶.

Seasonal forcing in transmission has been proposed as one determinant for the periodic patterns in other infectious diseases²⁷; however, Omori *et al.*²⁸ found that the seasonal forcing that occurs annually cannot generate the multi-year periodicity of MP incidence. They^{28,29} further proposed that the certain finite delay in the progression from immunity to the susceptible may provide an explanation to the occurrence of the cyclic epidemics of MP infections. More concretely, Omori *et al.*²⁸ show that “minor variation in the duration of immunity at the population level must be considered essential for the MP epidemic cycle because the MP cyclic incidence pattern did not replicate without it.” As shown in Fig. 3 of²⁸, this requires that the distribution for the duration of immunity should have a variance of around 0.63. Up to now no empirical data are available for estimating the distribution of the duration of MP immunity.

The MP incidence in England and Wales has declined⁷ following the widespread use of macrolides antibiotics since introduction in the late 1990s³⁰. Due to the emergence of macrolide-resistant strains, MP infections are of increasing public health interest^{31,32}. An understanding of the mechanisms by which recurrent epidemics of MP infection occur is urgently needed to enable control of future epidemics.

The two distinctive aspects of the MP epidemics: the prevalent serotype shifts among epidemics^{7,14,15,19,33} and cycling of MP incidence may be interconnected. This has been proposed before. It was argued^{16,17} that MP epidemics arise due to a change in the two main P1 types and variants of P1 sequences. Increased incidence of MP infection was observed to correlate with co-circulation of multiple strains within the population of England and

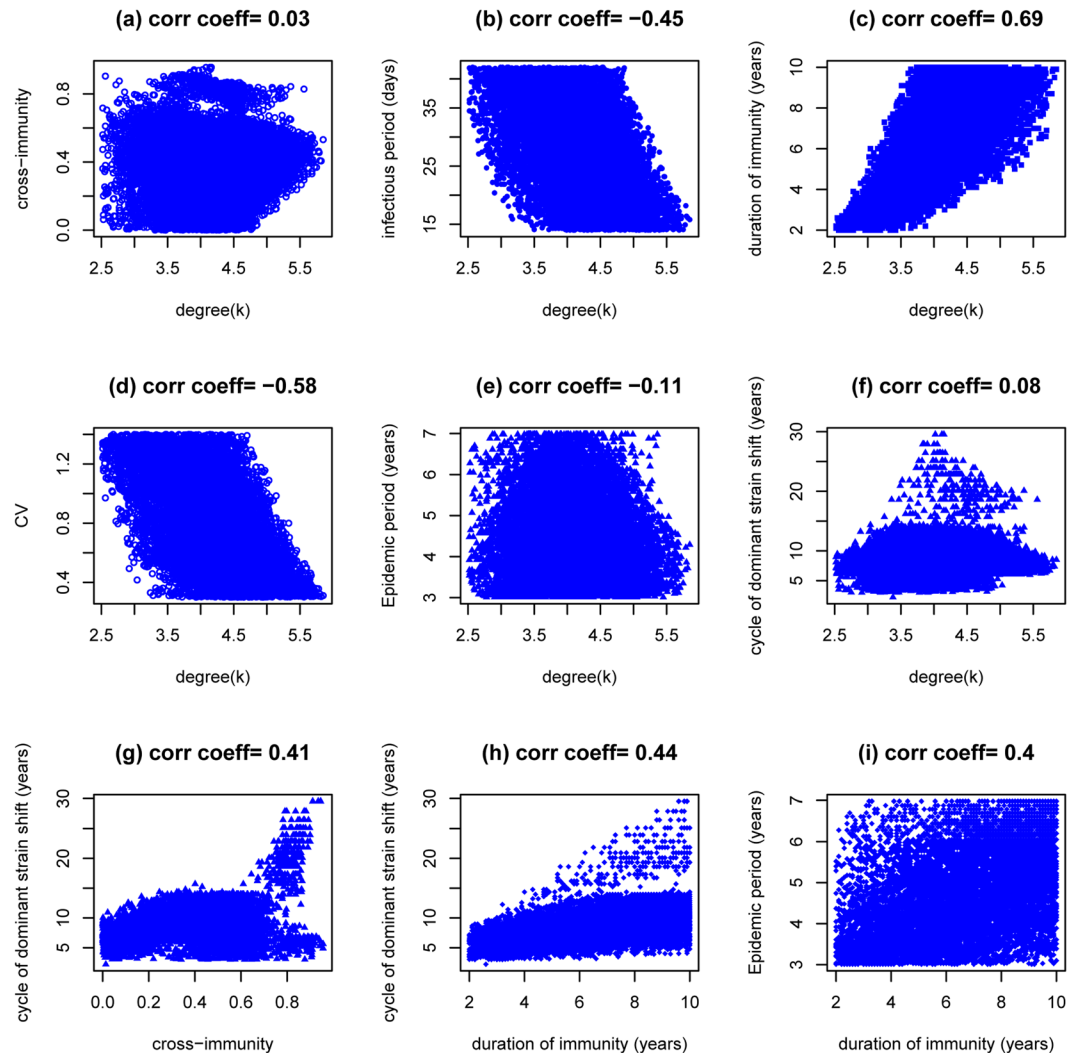


Figure 2. Features of LH sampling of model parameters and dynamic patterns of incidence. Here shown are dynamic patterns caused by two asynchronous strains under the condition of no interactions within the secondary infection (i.e., $\nu = \mu = 1$). The basic reproduction number $R_0 = 1.7$ is assumed.

Wales⁶. It was further speculated⁷ that dominant strain shift may be the cause of recurrent MP epidemics in view of the presence of multiple strains in observed increases of MP infection. Despite a lack of current data (due to limited focus on MP internationally and poor tools for detection and simultaneous strain discrimination) we speculate that serotype interactions such as synergistic associations and competition, in addition to the cross-immunity of differing P1 types, exist and play a possible role in the recurrent epidemics of MP infections. Previous transmission dynamics models (see the review of²⁸) neglected the following phenomena: co-circulation within human populations of multiple strains of MP and network structural contact patterns among people. Infection transmission depends on the contact rate as well as whom each individual contacts. Recent studies³⁴ showed that people do not mix randomly. For example, contact patterns between people may display the characteristics of scale-free networks³⁵ or small-world networks³⁶. An important parameter of a network is its degree, defined as the number of other individuals to which one is connected. A well-mixed network (i.e., the loose network) will have a high average degree while a less mixed network should have a small average. Realistic networks of contacts that are relevant to infectious diseases usually have a small average degree³⁷. On the contrary, the assumption of random mixing, in which every person is equally likely to contact any other person within the population^{27,38}, results in a very large degree.

Network structured models describe the transmission dynamics as in spatial transmission processes among connected groups and thus induce spatial correlation between infections. Letting infection spread on a homogeneous population with a fixed random network structure, Rozhnova and Nunes³⁹ illustrate that this spatial correlation within Susceptible-Infectious-Recovered-Susceptible models assists the generation of sustained cyclical epidemics. However, strong spatial correlation (i.e., strong network structure) was needed for the cycles to persist when they just considered the transmission dynamics of a single strain in a population. Considering a two

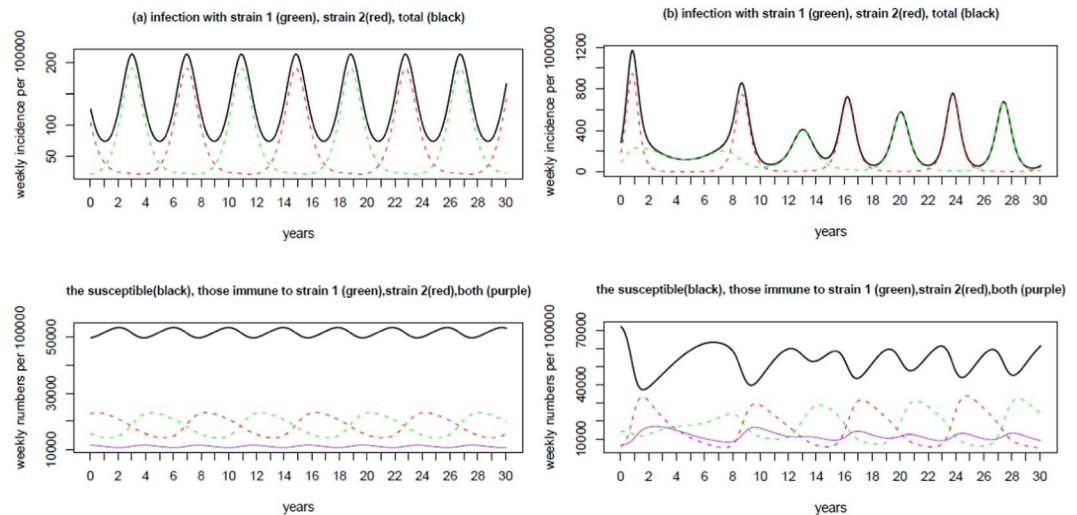


Figure 3. Two examples of epidemic curve from LH samples under the situation of no interaction within the secondary infection (i.e., $\nu = \mu = 1$). Panel (a) degree of contact (κ) = 5.69, cross-immunity (ψ) = 0.415, infectious period (D) = 17.2 days and duration of immunity (d) = 9.2 years, inter-epidemic period (EpiT) = 4 years, the coefficient of variance of the incidence time series (CV) = 0.51; panel (b) $\kappa = 3.34$, $\psi = 0.202$, $D = 37.7$ days, $d = 4.1$ years, and EpiT = 4 years, CV = 0.93. The legend is provided in the title to each figure. It is obvious that the proportion of individuals that are simultaneously immune to both strains is kept low in the two examples. The proportion of individuals immune to one strain alone is temporally highly anti-correlated with the proportion of these immune to the other strain alone, with the correlation coefficient -95% and -84% for panel (a) and panel (b) respectively.

strain version of the SIRS epidemic network model⁴⁰, the restriction on model parameters especially the degree of contacts is much relaxed. Recurrent epidemics were also predicted by models in a population which did not have a network structure, but in which people could be re-infected or co-infected with multiple strains⁴¹. Neither of these studies considered MP infection. In this study, we consider a two strain model as depicted in Fig. 1 and explore whether inclusion of both factors – a) competition between strains in a network-structured population and b) re-infection and interaction of secondary infection strain with primary infection strain – can explain the observed cycles in MP incidence.

Results

We study two specific scenarios in relation to the occurrence of MP epidemic cycles. First, assuming that two strains of MP interact only through the cross-immunity during the reinfection process we sought to explore how network-structured contacts alone can help build up the characteristics of MP epidemics. Secondly we assume that the primary infection can influence the infectivity and duration of infectiousness of a secondary infection, in addition to the cross-immunity. Under this situation, we examine how these strain interactions can help generate sustained recurrent epidemics and thus relax the requirement of network contacts for the build-up of MP epidemics.

Special situation I: network contact and cross-immunity alone. This is the situation where interaction parameters ν and μ are both set to 1 so strains don't interact within secondary infections. Preliminary sampling experiments show that under this situation the model system cannot generate recurrent epidemics when contact degree (κ) is greater than 6.0. 150,000 combinations of model parameters were sampled with $\nu = \mu = 1$ and κ ranging from 2.5 to 7.0. Only 11222 combinations generate characteristically recurrent epidemics of MP that are of asynchronous strains and their features are shown in Fig. 2. Other 214 combinations generate recurrent epidemics that are of synchronous strains (see Fig. S1). The results shown in Fig. 2a illustrate that reproducing the characteristically recurrent epidemics of MP is not possible unless the contact degree (κ) is less than 6.0. That is, without strain interaction within the secondary infection (i.e., $\nu = \mu = 1$), it requires strong network-mediated spatial correlation (c.f.^{39,40}) to enable MP epidemic cycling. When cross-immunity (ψ) is not extremely strong (shown in Fig. 2a), two strains asynchronously shift among epidemics. It is worth mentioning two extreme situations when $\psi = 0$ and 1. The former corresponds to the case of co-circulation of two independent strains. Figure 2a show that under this situation, recurrent epidemics of MP can be generated only when the contact degree (κ) is further less than 4.8. The latter corresponds to the case of single strain circulation^{39,40}. Recurrent epidemics can be produced when the contact degree (κ) is less than 4.2 (see Fig. S1) but the two strains completely synchronise.

For the model parameters that can generate recurrent epidemics of MP, the infectious period is negatively correlated while the duration of immunity is positively correlated with the degree of contact (Fig. 2b,c). This suggests that, all other parameters being equal: within a population of a relatively large contact degree, the infectious period will need to become shorter while the duration of immunity needs to become longer for recurrent

epidemics consistent with those observed to occur. Recurrent epidemics generated by a population of small contact degree (κ being just larger than 2.5) have a high coefficient of variation and show strong oscillations while those generated by the population of large contact degree (κ being just less than 6.0) have low coefficient of variation (Fig. 2d). Compared to other parameter combinations, both situations result in slightly shorter durations of recurrent epidemics and cycles of dominant strain shift (Fig. 2e,f), however. Cycles of dominant strain shift are positively associated with the presence of cross-immunity and the duration of immunity (Fig. 2g,h), whilst durations of recurrent epidemics are insensitive to these parameters (data not shown).

Two examples of the predicted recurrent epidemics are demonstrated in Fig. 3: one is a regular recurrent epidemic and the other irregular. To illustrate the possible mechanisms of oscillation in incidence and the shift of the dominant strain, we also plot the changes in the susceptible individuals, and the individuals that are immune to strain 1 alone, and to strain 2 alone, and to both strains together. In Fig. 3a MP epidemics occur regularly with epidemic period of exactly 4 years and the coefficient of variance (CV) of the incidence time series is 0.51. Two strains alternate the dominancy symmetrically from one epidemic to another: when one strain is dominant the other strain remains at extremely low activities. That is, each separate epidemic is mainly caused by one strain. In Fig. 3b the duration of recurrent epidemics ranges from 3 years to 5 years with an average of four years. The average CV is 0.93, indicating a strong oscillation comparing to example shown in Fig. 3a. The epidemics also vary in the total number of infections. During each epidemic, infections can be due to either mainly one strain or two strains simultaneously.

Comparisons of the upper and bottom graphs in each panel of Fig. 3 show that infections oscillate following the changes in proportion of the population that is susceptible. The shift of the dominant strain during the oscillating epidemics is due to changes in the proportion of the population that is immune to different strains: The incidence of one strain will increase when the proportion of people immune to it is low; at the same time the incidence of the other strain will decrease because of the relatively high proportion of the population that is immune. This observation seems to support the hypothesis that a decline in immunity or an increase of the immunologically naïve population may result in the 4-year cycle of epidemic periods⁶.

Special situation II: network contact and strain interactions via cross-immunity during re-infection and interactions within secondary infection.

Preliminary sampling experiments indicate that the number of parameter combinations that can generate MP recurrent epidemics decreases quickly as the degree of contacts increases. When the degree of contacts (κ) exceeds 15, the combinations of model parameters for MP recurrent epidemics become extremely rare. To save the computational time, the model parameter values were sampled by dividing them into groups by the ranges of κ : 2.5–6, 6–8, 8–10, 10–12, 12–14, 14–16, 16–17 with respective sampling sizes 200,000, 200,000, 250,000, 250,000, 250,000, 250,000, 500,000. We obtained 23534 combinations of model parameter values that generate characteristically recurrent epidemics of MP that are of asynchronous strains; the maximum of κ is 16.2 (see Fig. 4a). Other 4426 combinations generate recurrent epidemics that are of synchronous strains (see Fig. S2).

The results shown in Fig. 4 indicate that compared to the above situation (I), allowing for strain interactions within secondary infections can lead to the characteristically recurrent epidemics of MP even in a population that has little network mediated spatial correlation. The maximum degree of contacts (16.2) is much larger than the maximum value of 5.9 that was required for situation I. The distributed patterns in Fig. 4a–c suggest that there is a critical threshold in the degree of contacts (denoted by κ_{ac} for asynchronous strain recurrent epidemics thereafter in the paper) separating the mechanisms by which recurrent epidemics consistent with those seen for MP occur. For the parameter values given in Fig. 4, $\kappa_{ac} = 6.0$. For populations that are of contact degree $\kappa < \kappa_{ac}$, recurrent epidemics occur because of the spatial correlation induced by strong network structure whilst for the populations of relatively loose network structure ($\kappa > \kappa_{ac}$), they occur because of the combination of spatial correlation and strain interactions. We refer them as mechanism 1 and mechanism 2 respectively. It is clearly shown in Fig. 4a that although cross-immunity can be any level from 0 to 1 under mechanism 1, only weak cross-immunity levels (< 0.4) are required under mechanism 2.

As shown above (special situation I), when $\kappa < \kappa_{ac}$ (mechanism 1), characteristically MP recurrent epidemics are readily generated irrespective of whether the primary infection affects the secondary infection. When strain interactions are present, complicated epidemics can be generated. Under the loose network structure ($\kappa > \kappa_{ac}$) (mechanism 2), MP recurrent epidemics can be produced only when the primary infection enhances the infectivity ($\nu > 1$) (Fig. 4b) and prolongs the infectious period ($\mu > 1$) (Fig. 4c) of the secondary infection. Conversely, if strain interactions diminish the transmissibility of a secondary infection, shifts in the dominant strain in epidemics cannot occur. As in special situation I, relatively short infectious periods of MP at populations of high contact degree are required while relatively longer durations of immunity are needed to generate the recurrent MP epidemics (Fig. 4d,e). The conditions for the emergence of synchronous strains are different and are shown in Fig. S2.

Dynamic patterns including the shape of oscillations and durations of recurrent epidemics and cycle of dominant strain shift are shown in Fig. 4f–l. Figure 4f shows that the shape of the epidemic curve (i.e., CV of incidence along the time) is positively correlated with κ when $\kappa > \kappa_{ac}$, while CV decreases with κ when $\kappa \leq \kappa_{ac}$. This suggests that within a looser networked population, the oscillation in incidence tends to become stronger. However, CV is not sensitive to the other model parameters (data not shown).

The cycle of dominant strain shift ranges from 3 to 30 years (Fig. 4g), which covers the observational ranges: 10–16 years^{14,15}. Figure 4j indicates that when $\kappa < \kappa_{ac}$, the cycle of dominant strain shift can be 1–6 times the duration of recurrent epidemics; when $\kappa > \kappa_{ac}$, the cycle of dominant strain shift approximates the epidemic cycle. Both the duration of recurrent epidemics and cycle of dominant strain shift are positively associated with the duration of immunity, especially under mechanism 2 (Fig. 4h,i). Under mechanism 2, they are weakly and positively associated with the infectious period (Fig. 4i,k). Otherwise, they are insensitive to other parameters (see Fig. 4g for the relationship between cycle of dominant strain shift and degree of contacts).

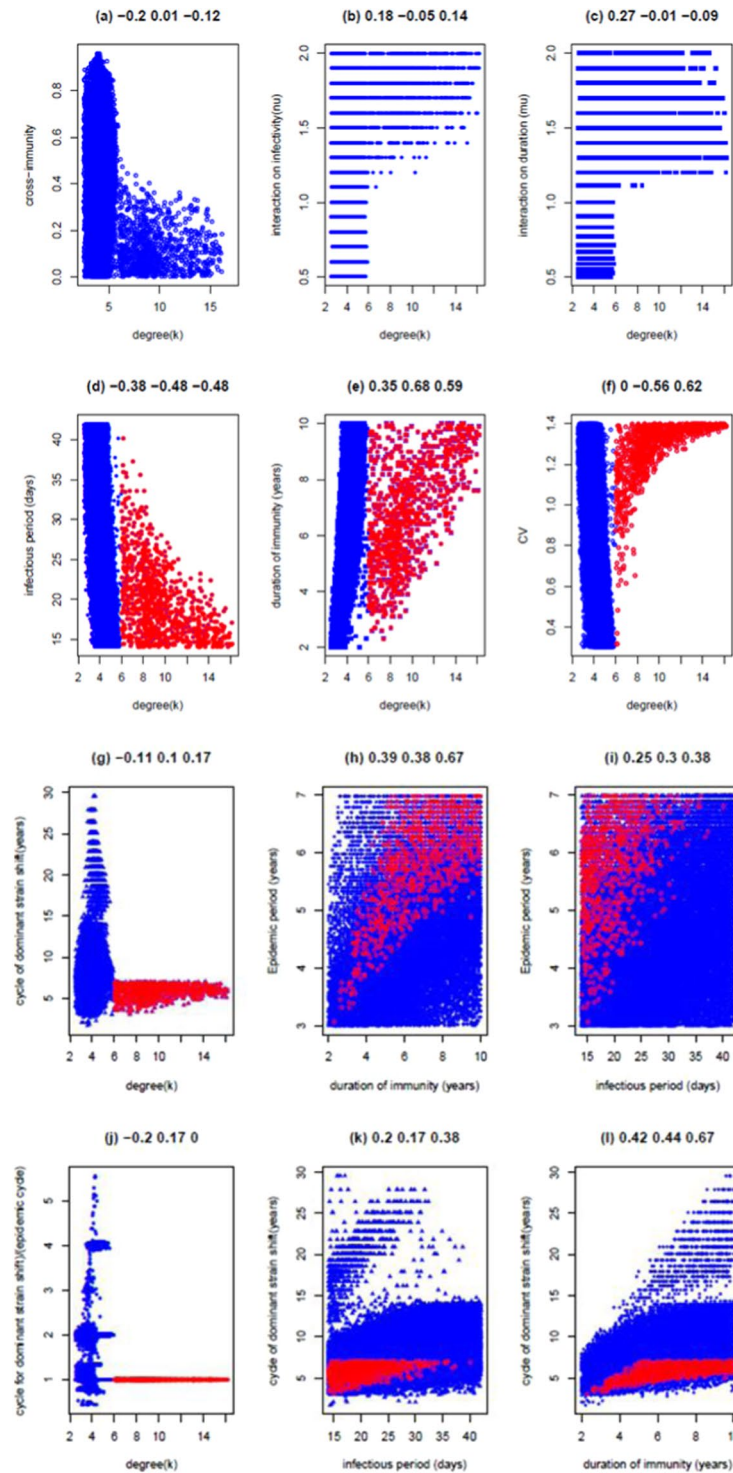


Figure 4. Features of LH samplings of model parameters and dynamic patterns of incidence. Here shown are dynamic patterns caused by two asynchronous strains that interact within the secondary infection. The basic reproduction number $R_0 = 1.7$ is assumed. Panel (a) shows the maximum degree of contacts is 16.2 while panels (b) and (c) show that the critical degree for the asynchronous strains is $\kappa_{ac} = 6.0$. In panels (d–l) the blue points represent the parameter values when contact degrees $\kappa \leq \kappa_{ac}$ and the red points for those with contact degrees $\kappa > \kappa_{ac}$. The three values above each panel represent the correlation coefficients between the two variables for all values, the values when $\kappa \leq \kappa_{ac}$, and the values when $\kappa > \kappa_{ac}$.

Four typically recurrent epidemic examples are illustrated in Fig. 5. They show different oscillation patterns. In panel (a) two strains are of comparable activity levels with the strain that starts early dominating the epidemics; strain dominance alternates regularly among epidemics cycle and oscillate with the same period of four years. In

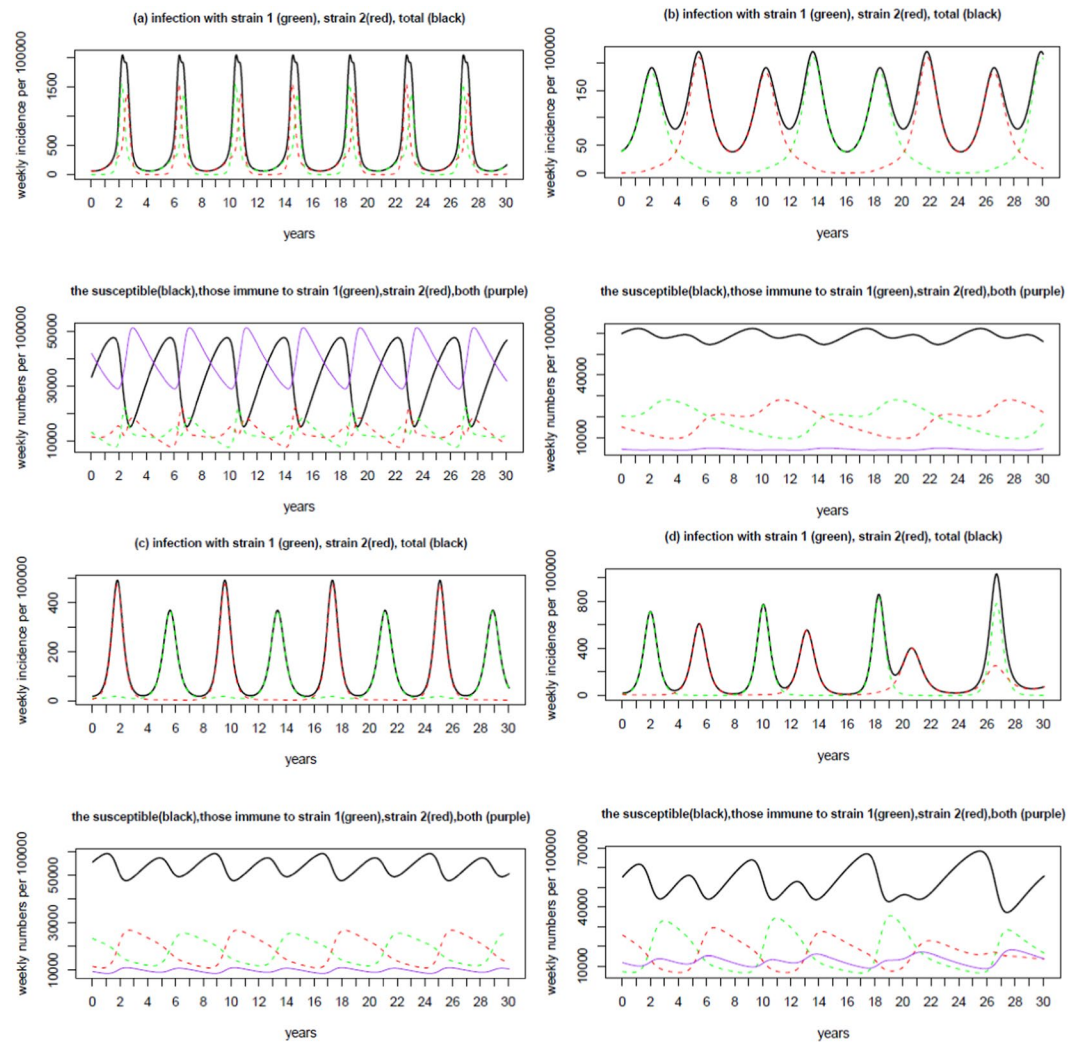


Figure 5. Four examples of epidemic curve from LH samples that generate MP recurrent epidemics with two asynchronous strains that interact within secondary infections. Panel (a) degree of contact (κ) = 8.8, cross-immunity (ψ) = 0.01, infectious period (D) = 15.4 days, duration of immunity (d) = 4.3 years, relative infectiousness of secondary infection (ν) = 1.8, relative infectious period of secondary infection (μ) = 1.43, and inter-epidemic period (EpiT) = 4.0 years, the coefficient of variance of the incidence time series (CV) = 1.16; (b) κ = 5.1, ψ = 0.70, D = 14.4 days, d = 8.6 years, ν = 1.4, μ = 0.59, and EpiT = 4.0 years, CV = 0.48; (c) κ = 5.0, ψ = 0.42, D = 14.1 days, d = 8.8 years, ν = 1.2, μ = 0.77, and EpiT = 4.0 years, CV = 0.90; (d) κ = 3.67, ψ = 0.19, D = 32.6 days, d = 5.4 years, ν = 1.6, μ = 0.63, and EpiT = 4.0 years, CV = 0.60. The legend is provided in the title to each figure.

panel (b) two strains shift dominance with each strain dominating two epidemics consecutively before switching strain dominance; the two consecutive epidemics are mainly activated by the dominant strain while the other strain remains at very low activity. In panel (c) although epidemics take place regularly, recurrent epidemics consist of two different epidemics: one with high peak and narrow active period, the other with lower peak but wide active period; two strains alternate their dominance accordingly. In panel (d) two strains alternate with irregular peaks and total incidence within each epidemic. The diverse patterns may mimic real observations in MP epidemics^{7,14}.

Comparing the levels of infection and of immunity can shed light on the underlying mechanisms of recurrent epidemics. It is obvious from the Fig. 5 that the proportion of individuals that are simultaneously immune to both strains is kept low except for panel (a) where it oscillates within a wide range and anti-correlates with the proportion susceptible. The proportion of individuals immune to one strain is temporally highly anti-correlated with the proportion of those immune to the other strain, with absolute correlation coefficients $> 80\%$ for panels (b), (c) and (d); while for panel (a) they are weakly correlated. This difference reflects their different levels of cross-immunity. Panels (a) and (b) show predictions obtained for a situation in which the primary infection strain increases the infectivity of secondary strain ($\nu > 1$). In panel (a) although the dominant strain shifts between epidemics, the difference in the proportion immune or susceptible to the dominant and non-dominant strains is

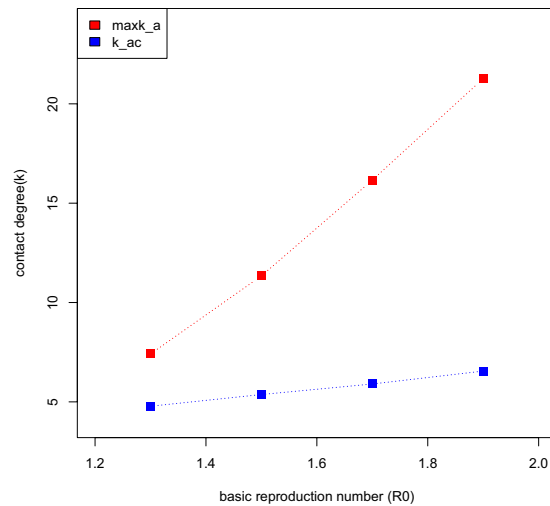


Figure 6. Critical threshold degrees of contacts under different transmissibility.

small. In panel (b), one strain is dominant while the other remains at a very low incidence, which continues over further epidemics even if the proportion of individuals that are immune to the strain exceeds the proportion that is immune to the other strain. The dominance only changes when the difference in immunity to a given strain increases substantially. So under this situation, every strain dominates continuously over two epidemics before the strain dominance switches. As found for situation I, oscillations in the infection incidence follow changes in the proportion of the population that is immune and susceptible. The change of dominant strain during the recurrent epidemics is due to the exchange in immunity to different strain: the increase in infection activity of one strain follows the relatively low immunity to the strain⁶.

The findings assuming values for R_0 of 1.3 and 1.5 are similar to those obtained assuming a value of 1.7, although the critical threshold value in the degree of contacts differs (see Fig. 6). Under a low value for R_0 of 1.3, for example, the critical degree of contact decreases to $\kappa_{ac} = 4.7$ and the maximum contact degree decreases to 7.4.

Discussion

In this study we demonstrate that spatial correlation mediated by contact network of human population and positive strain interactions within secondary infection work cooperatively to drive MP infection incidence into recurrent epidemics occurring every three to seven years with dominant strains shifting among epidemics.

Accounting for realistic host population structure in infectious disease modelling is important. It has been recognised that it is necessary to take true network contacts among human populations to explain the observed dissemination patterns of infectious disease (e.g.⁴²). The results shown in Fig. 2 where no strain interaction is assumed illustrate that contact networks of degree less than 6 are required for the recurrent epidemics that occur every 3–7 years with alternation of dominant strain. A key property of a network is its degree distribution. For community, school and hospital networks, empirical studies suggest that the average degree is 6.5³⁷. Other studies show higher average degrees; For example, it was estimated to be about 8.3 within social encounter networks in Great Britain⁴³ and 12.5 among Hong Kong residents, a population inhabiting one of the highest density locations in the world⁴⁴. If this empirical information reflected the true characteristics of human contact networks, then the network model without strain interaction within secondary infection could not be a candidate mechanism for MP recurrent epidemics.

Interaction between different strains during re-infection (i.e. cross-immunity) is well known and has attracted much effort to study and measure it. Once an individual is re-infected by another strain, are there any interactions between the primary infection strain and the secondary infection strain within the secondary infection? Surely recovery from primary infection will not leave immunocompetent host individuals naïve. It is theoretically reasonable to argue that the non-naïve individuals would have other changes which might in some ways alter the secondary infection by other strains (see⁴¹ for more general reasoning). For example, primary infection with one strain of dengue virus can enhance transmission of subsequent infection with another strain⁴⁵. This occurs through a mechanism called “intrinsic antibody-dependent infection enhancement” and this phenomenon might be involved in many viral and bacterial infections⁴⁶.

As to MP, it parasitizes the respiratory tract epithelium of humans, the primary infection with one strain can, for example, damage the airway⁴⁷, which could then alter the ecological niche of the secondary infection strain. Further, the primary infections of the upper or lower respiratory tract can be followed by extrapulmonary complications⁴⁸. As far as the transmission dynamics are concerned, the modifications in the non-naïve individuals might change the infectivity and duration of secondary infection. To our knowledge, we have not found any clear empirical data for these interactions among strains of MP, although this may reflect a lack of research into MP pathogenicity. In principle, prior exposure of an individual to a strain could have no effect or either decrease or

increase the individual's ability to clear an infection with a differing strain, with potential to increase or reduce the overall transmission. The theoretical analysis in this study shows that only positive strain interactions increase the infectiousness of a secondary infection to facilitate the generation of recurrent MP epidemics. It is expected that the experimental observations and measurement of strain interactions within secondary infections will provide vital proof to support or disprove the combination of network mediated spatial correlations and strain interactions within secondary infections as a determinant of epidemic recycling of MP.

We found that there is a positive association between the durations of recurrent epidemics and the duration of immunity (Fig. 4h). This finding is consistent with that of⁴⁹ but differ from that of²⁸. Further, the cycle of dominant strain shift also shows positive correlations with the duration of immunity (Fig. 4l). These positive correlations become stronger under the mechanism whereby recurrent epidemics are generated by the combination of network mediated spatial correlations and strain interactions within secondary infection. Under this mechanism, both cycles are intermediately linked with the infectious period (Fig. 4i,k). In contrast to the observations of^{27,28}, the cycle of the dominant strain shift is insensitive to cross-immunity.

Despite our simplifying assumptions, the network model of the transmission dynamics of two strains presented here remains complicated. The distributions of both infectious period and the duration of immunity are implicitly assumed to be exponential. Omori *et al.*²⁸ suggested that assuming that the duration of immunity follows a distribution with a variance about 0.63, which is much smaller than that of the exponential distribution, models that assume that people mix randomly can produce the periodicity of MP recurrent epidemics. Can the reduced variation in the distribution of the duration of immunity help build up the recurrent epidemics in our network model? To see this, we construct a $SI_1I_2R_1R_2J_1J_2RR$ network model by separating recovery stages into two equal parts. Therein the immunity period follows a gamma distribution of shape parameter = 2. This model has 11 nodes and is described by 65 differential equations. Nonetheless, the simulations (data not shown) show that this more complicated model does not give any noticeably different results. It is a technical challenge in our network model to generate gamma distributed duration of immunity that is comparable to that required in²⁸.

We modelled human population contact structure as a static, unweighted network wherein each individual has an equal number of links with other people (i.e. regular random network), although it is recognised that in reality, the number of contacts made by each person differs between individuals, varying by age and other population groups (e.g.^{43,44,50,51}). One previous study⁵⁰ showed that the regular random network may offer a reasonable approximation to highly connected networks that approximate human contact patterns. Whilst the heterogeneity in contact networks is likely to affect the occurrence and size of outbreaks, its effect on the long-term dynamics of MP is unclear, given the low infectiousness of MP. For example, studies⁴³ have postulated that, unless an infection is highly infectious, such as norovirus or some haemorrhagic fevers, the total time spent with others is more important for its transmission than the number of contacts made. This follows if, as is plausible, the duration of the contact decreases as the number of contacts made per person increases. Elucidating these issues is beyond the scope of this paper and will be studied in future work.

In this study we constrained both the interaction parameters ν and μ , which describe the effects of primary infection strain on the infectivity and duration of infection by secondary strain respectively, at the ranges from 0.5 to 2.0. If we had widened their ranges, requirement for the limited contact degree is further reduced. This is in agreement with the previous studies⁴¹: under strong strain interactions alone, epidemic cycling becomes possible even under assumptions of the homogenous mixing.

To search for the model parameter combinations that generate the recurrent epidemics of MP, Latin hypercube sampling⁵² was employed. As a statistical method for generating a near-random sample of parameter values from a multidimensional distribution, it has a short come: we do not know how many parameter combinations are adequate to guarantee that a full range of all possible outcomes are explored. This difficulty has been mentioned in Results section. Another method of sampling multi-dimensional parameters recommended by⁵³ can avoid this difficulty through total sensitivity analysis⁵⁴. This surely will merit using in the studies like ours.

In conclusion, we have illustrated that multiple strains that co-circulate within a network structured population and interact positively as secondary infections with primary infections generate the MP epidemics of 3–7 year interval and alternating dominant strains. This model supports the theory that epidemic shifts in MP may be attributed to population immunity not only to the immunogenic strain in question, but also with the influence of cross protection and other enhanced effects from the second strain type and that transmission via patient networks within the population combine to produce MP epidemic cycles. Though the strain interactions within a secondary infection are theoretically possible, currently no reliable evidence exists to suggest whether either a positive or negative strain interaction occurs. We hope this study can encourage experimental studies to detect and measure interactions between strains of MP. This will benefit our understanding of MP and provide crucial information for us to predict and thus control its recurrence.

Models and Methods

General structure of the model. We consider a Susceptible-Infectious-Recovered-Susceptible (SIRS) transmission dynamics model of two strains in view of the fact that isolates and strains of MP can be classified into two main types: P1 type 1 and P1 type 2^{14,16,20,21}. Within the SIRS model, a population of size N is modelled as a network in which every individual randomly contacts a fixed number (κ) of other individuals, and is classified into eight compartments (Fig. 1), namely those who are susceptible to infection with any strain (S), those who are infected and infectious with strain 1 or 2 (I_1 and I_2), those who have recovered from infection with a given strain and are susceptible to infection with the other strain (R_1 and R_2), those who are infected and infectious with

strain 1 or 2 after recovering from previous infection (J_1 and J_2) and those who are immune to infection with both strains (R). We refer to people in the I_1 and I_2 compartments as those with “primary infection”, and to people in the J_1 and J_2 compartments as those with “secondary infection”. Individuals are denoted by nodes and contacts between individuals by edges.

The epidemic dynamics is determined by the following transmission and transition processes. Susceptible nodes (S) become infected with strain i , $i = \{1, 2\}$, at rate λ through an edge with a node of primary infection I_i , or at rate $\nu\lambda$ through an edge with a node of secondary infection J_i . Here parameter λ represents the constant transmission rate and parameter ν is the infectiousness of a secondary infection relative to a primary infection. Primarily infected nodes (I_i) stay infectious on average for D days before becoming fully immune (R_i) to the infecting strain i and partially so to the other strain. Recovered individuals (R_i) stay immune for an average of d days before becoming susceptible again, or becoming secondarily infected at rate $(1-\psi)\lambda$ through an edge linked with a node of infection (I_{3-i} or J_{3-i}) to become secondarily infected J_{3-i} , $i = \{1, 2\}$. Here ψ reflects the reduction in susceptibility due to previous exposure to the other strain (i.e., cross-immunity). Nodes of secondary infection J_i , $i = \{1, 2\}$ stay infectious for an average of μD days before becoming fully immune to all strains (i.e., R). Here parameter μ defines the effect of having experienced primary infection on the duration of the secondary infection. Nodes of R stay fully immune for an average of d days before becoming susceptible again. Therefore naïve individuals are recruited into the population through birth and loss of immunity. These transitions and transmissions are defined according to the pairs or triplets involved in the process^{39,40,55}. For simplicity we ignore clustering in the network (c.f. ^{37,40,55}).

Model equations. Similar to⁵⁵, the proportions of people in eight compartments are represented by $[S]$, $[I_1]$, $[I_2]$, $[R_1]$, $[R_2]$, $[J_1]$, $[J_2]$, and $[R]$. Because of the constant population size (i.e., the constant number of nodes), $[S] + [I_1] + [I_2] + [R_1] + [R_2] + [J_1] + [J_2] + [R] = 1$. There are $(8 \times 7)/2 = 28$ heterogeneous pairs within the network in which the two nodes of a pair are of different states. The proportion of the population that is in a pair ($[XY]$) is defined as

$$[XY] \approx \frac{n_{XY}}{\kappa N} \quad (1)$$

Here n_{XY} is the number of pairs within the population. The number of homogenous pairs can be found from these equations for heterogeneous pairs: e.g., $[RR] = (1 - \sum_{Y \neq R} [Y]) - \sum_{X \neq R} [XR]$. The state of the model system is defined by eight integers of nodes and 28 integers of heterogeneous pairs. To focus on the impact of spatial correlation mediated by network structure and interactions between strains, two strains are simply assumed to be antigenically indistinguishable within linked patients although they can actually be different⁵⁶.

Infection transmission occurs through pair-link and the change of pairs is determined by the triples⁴⁰. To close the model system, the proportion, $[XYZ]$, of the triple XYZ with node Y having contacts with both X and Z is approximated in terms of the proportion of pairs as in⁵⁵,

$$[XYZ] \approx \frac{k-1}{k} \frac{[XY][YZ]}{[Y]} \quad (2)$$

The flow chart of the epidemic model is shown in Fig. 1. The model is described by a set of $28 + 7 = 35$ differential equations as,

Equations describing the time changes in 7 nodes

$$\begin{aligned} \frac{d}{dt}[S] &= \frac{1}{L}(1 - [S]) + \frac{1}{d}([R] + [R_1] + [R_2]) - \kappa(F_1 + F_2) \\ \frac{d}{dt}[I_1] &= -\left(\frac{1}{L} + \frac{1}{D}\right)[I_1] + \kappa F_1 \\ \frac{d}{dt}[I_2] &= -\left(\frac{1}{L} + \frac{1}{D}\right)[I_2] + \kappa F_2 \\ \frac{d}{dt}[R_1] &= -\left(\frac{1}{L} + \frac{1}{d}\right)[R_1] + \frac{1}{D}[I_1] - \kappa G_2 \\ \frac{d}{dt}[R_2] &= -\left(\frac{1}{L} + \frac{1}{d}\right)[R_2] + \frac{1}{D}[I_2] - \kappa G_1 \\ \frac{d}{dt}[J_1] &= -\left(\frac{1}{L} + \frac{1}{\mu D}\right)[J_1] + \kappa G_1 \\ \frac{d}{dt}[J_2] &= -\left(\frac{1}{L} + \frac{1}{\mu D}\right)[J_2] + \kappa G_2 \end{aligned} \quad (3a)$$

Equations describing the time changes of 28 pairs

$$\begin{aligned}
\frac{d}{dt}[SI_1] &= \frac{1}{a}([I_1R] + [I_1R_2] + [I_1R_1]) - \left(\frac{1}{D} + \lambda\right)[SI_1] \\
&\quad + (\kappa - 1)\left\{\frac{([SS] - [SI_1])F_1 - [SI_1]F_2}{[S]}\right\} + \frac{1}{L}([I_1] - 2[SI_1]) \\
\frac{d}{dt}[SI_2] &= \frac{1}{a}([I_2R] + [I_2R_2] + [I_2R_1]) - \left(\frac{1}{D} + \lambda\right)[SI_2] \\
&\quad + (\kappa - 1)\left\{\frac{([SS] - [SI_2])F_2 - [SI_2]F_1}{[S]}\right\} + \frac{1}{L}([I_2] - 2[SI_2]) \\
\frac{d}{dt}[SR_1] &= \frac{1}{a}([R_1R] + ([R_1R_1] - [SR_1]) + [R_1R_2]) + \frac{1}{D}[SI_1] \\
&\quad - (\kappa - 1)\left\{\frac{F_1 + F_2}{[S]} + \frac{G_2}{[R_1]}\right\}[SR_1] + \frac{1}{L}([R_1] - 2[SR_1]) \\
\frac{d}{dt}[SR_2] &= \frac{1}{a}([R_2R] + ([R_2R_2] - [SR_2]) + [R_1R_2]) + \frac{1}{D}[SI_2] \\
&\quad - (\kappa - 1)\left\{\frac{F_1 + F_2}{[S]} + \frac{G_1}{[R_2]}\right\}[SR_2] + \frac{1}{L}([R_2] - 2[SR_2]) \\
\frac{d}{dt}[SR] &= \frac{1}{a}([RR] - [SR]) + [R_2R] + [R_1R] + \frac{1}{\mu D}([SJ_1] + [SJ_2]) \\
&\quad - \frac{(\kappa - 1)[SR][F_1 + F_2]}{[S]} + \frac{1}{L}([R] - 2[SR]) \\
\frac{d}{dt}[SJ_1] &= \frac{1}{a}([R_1J_1] + [R_2J_1] + [R_1J_1]) - \left(\frac{1}{\mu D} + \nu\lambda\right)[SJ_1] \\
&\quad - (\kappa - 1)\left\{\frac{[SJ_1](F_1 + F_2)}{[S]} - \frac{[SR_2]G_1}{[R_2]}\right\} + \frac{1}{L}([J_1] - 2[SJ_1]) \\
\frac{d}{dt}[SJ_2] &= \frac{1}{a}([R_1J_2] + [R_2J_2] + [R_1J_2]) - \left(\frac{1}{\mu D} + \nu\lambda\right)[SJ_2] \\
&\quad - (\kappa - 1)\left\{\frac{[SJ_2](F_1 + F_2)}{[S]} - \frac{[SR_1]G_2}{[R_1]}\right\} + \frac{1}{L}([J_2] - 2[SJ_2]) \\
\frac{d}{dt}[II_2] &= -2\left(\frac{1}{L} + \frac{1}{D}\right)[II_2] + (\kappa - 1)\left\{\frac{[SI_2]F_1}{[S]} + \frac{[SI_1]F_2}{[S]}\right\} \\
\frac{d}{dt}[I_1R_1] &= -\left(\frac{2}{L} + \frac{1}{a}\right)[I_1R_1] + \frac{1}{D}([I_1] - [I_1R_1]) + (\kappa - 1) \\
&\quad \times \left\{\frac{[SR_1]F_1}{[S]} - [I_1R_1]\frac{G_2}{[R_1]}\right\} \\
\frac{dp}{dt}[I_1R_2] &= -\left\{\frac{2}{L} + \frac{1}{a} + \frac{1}{D} + \lambda(1 - \psi)\right\}[I_1R_2] + \frac{1}{D}[I_1I_2] + (\kappa - 1) \\
&\quad \times \left\{\frac{[SR_2]F_1}{[S]} - [I_1R_2]\frac{G_1}{[R_2]}\right\} \\
\frac{d}{dt}[I_1R] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{a}\right)[I_1R] + \frac{1}{\mu D}([I_1] + [I_1I_2]) + (\kappa - 1)\frac{[SR]F_1}{[S]} \\
\frac{d}{dt}[I_1J_1] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{\mu D}\right)[I_1J_1] + \nu\lambda[SJ_1] + \lambda(1 - \psi)[I_1R_2] + (\kappa - 1) \\
&\quad \times \left\{\frac{[SJ_1]F_1}{[S]} + \frac{[I_1R_2]G_1}{[R_2]}\right\} \\
\frac{d}{dt}[I_1I_2] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{\mu D}\right)[I_1I_2] + (\kappa - 1)\left\{\frac{[SJ_2]F_1}{[S]} + \frac{[I_1R_1]G_2}{[R_1]}\right\} \\
\frac{d}{dt}[I_2R_1] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{a} + \lambda(1 - \psi)\right)[I_2R_1] + \frac{1}{D}[I_1I_2] + (\kappa - 1) \\
&\quad \times \left\{\frac{[SR_1]F_2}{[S]} - [I_2R_1]\frac{G_2}{[R_1]}\right\} \\
\frac{d}{dt}[I_2R_2] &= -\left(\frac{2}{L} + \frac{1}{a}\right)[I_2R_2] + \frac{1}{D}([I_2] - [I_2R_2]) + (\kappa - 1) \\
&\quad \times \left\{\frac{[SR_2]F_2}{[S]} - [I_2R_2]\frac{G_1}{[R_2]}\right\}
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}[I_2R] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{d}\right)[I_2R] + \frac{1}{\mu D}([I_2J_1] + [I_2J_2]) + (\kappa - 1)\frac{[SR]F_2}{[S]} \\
\frac{d}{dt}[I_2J_1] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{\mu D}\right)[I_2J_1] + (\kappa - 1)\left\{\frac{[SJ_1]F_2}{[S]} + \frac{[I_2R_2]G_1}{[R_2]}\right\} \\
\frac{d}{dt}[I_2J_2] &= -\left(\frac{2}{L} + \frac{1}{D} + \frac{1}{\mu D}\right)[I_2J_2] + \nu\lambda[SJ_2] + \lambda(1 - \psi)[I_2R_1] + (\kappa - 1) \\
&\quad \times \left\{\frac{[SJ_2]F_2}{[S]} + \frac{[I_2R_1]G_2}{[R_1]}\right\} \\
\frac{d}{dt}[R_1R_2] &= -2\left(\frac{1}{L} + \frac{1}{d}\right)[R_1R_2] + \frac{1}{D}([I_1R_2] + [I_2R_1]) - (\kappa - 1)[R_1R_2]\left\{\frac{G_2}{[R_1]} + \frac{G_1}{[R_2]}\right\} \\
\frac{d}{dt}[R_1R] &= -2\left(\frac{1}{L} + \frac{1}{d}\right)[R_1R] + \frac{1}{D}[I_1R] + \frac{1}{\mu D}([R_1J_1] + [R_1J_2]) - \frac{(\kappa - 1)[R_1R]G_2}{[R_1]} \\
\frac{d}{dt}[R_1J_1] &= -\left(\frac{2}{L} + \frac{1}{d} + \frac{1}{\mu D}\right)[R_1J_1] + \frac{1}{D}[I_1J_1] - (\kappa - 1)\left\{\frac{[R_1J_1]G_2}{[R_1]} - \frac{[R_1R_2]G_1}{[R_2]}\right\} \\
\frac{d}{dt}[R_1J_2] &= -\left\{\frac{2}{L} + \frac{1}{d} + \frac{1}{\mu D} + \nu(1 - \psi)\lambda\right\}[R_1J_2] + \frac{1}{D}[I_1J_2] \\
&\quad + \frac{(\kappa - 1)([R_1R_1] - [R_1J_2])G_2}{[R_1]} \\
\frac{d}{dt}[R_2R] &= -2\left(\frac{1}{L} + \frac{1}{d}\right)[R_2R] + \frac{1}{D}[I_2R] + \frac{1}{\mu D}([R_2J_1] + [R_2J_2]) - \frac{(\kappa - 1)[R_2R]G_1}{[R_2]} \\
\frac{d}{dt}[R_2J_1] &= -\left\{\frac{2}{L} + \frac{1}{d} + \frac{1}{\mu D} + \nu(1 - \psi)\lambda\right\}[R_2J_1] + \frac{1}{D}[I_2J_1] \\
&\quad + \frac{(\kappa - 1)([R_2R_2] - [R_2J_1])G_1}{[R_2]} \\
\frac{d}{dt}[R_2J_2] &= -\left(\frac{2}{L} + \frac{1}{d} + \frac{1}{\mu D}\right)[R_2J_2] + \frac{1}{D}[I_2J_2] + (\kappa - 1)\left\{\frac{[R_1R_2]G_2}{[R_1]} - \frac{[R_2J_2]G_1}{[R_2]}\right\} \\
\frac{d}{dt}[RJ_1] &= -\left(\frac{2}{L} + \frac{1}{d} + \frac{1}{\mu D}\right)[RJ_1] + \frac{1}{\mu D}([J_1J_1] + [J_1J_2]) + \frac{(\kappa - 1)[R_2R]G_1}{[R_2]} \\
\frac{d}{dt}[RJ_2] &= -\left(\frac{2}{L} + \frac{1}{d} + \frac{1}{\mu D}\right)[RJ_2] + \frac{1}{\mu D}([J_2J_2] + [J_1J_2]) + \frac{(\kappa - 1)[R_1R]G_2}{[R_1]} \\
\frac{d}{dt}[JJ_2] &= -2\left(\frac{1}{L} + \frac{1}{\mu D}\right)[JJ_2] + (\kappa - 1)\left\{\frac{[R_2J_2]G_1}{[R_2]} + \frac{[RJ_1]G_2}{[R_1]}\right\}
\end{aligned} \tag{3b}$$

In the above equations, the different forces of infection are

$$\begin{aligned}
\text{Strain 1 infects } S \text{ individuals: } F_1 &= \lambda([SI_1] + \nu[SJ_1]) \\
\text{Strain 2 infects } S \text{ individuals: } F_2 &= \lambda([SI_2] + \nu[SJ_2]) \\
\text{Strain 1 infects } R_2 \text{ individuals: } G_1 &= \lambda(1 - \psi)([I_1R_2] + \nu[J_1R_2]) \\
\text{Strain 2 infects } R_1 \text{ individuals: } G_2 &= \lambda(1 - \psi)([I_2R_1] + \nu[R_1J_2])
\end{aligned} \tag{4}$$

The model parameters are defined in Table 1. Compared with the model presented in⁴⁰, here we introduce two interaction parameters to define the effects of experiencing primary infection on infectivity (ν) and the duration (μ) of a secondary infection. The complexity of the two strain network dynamics allows us to investigate the combined effects of strain interactions (cross-immunity during re-infection and effects of the primary infection on a secondary infection) on dynamic patterns of endemic infectious diseases, along with spatial competition embedded within the random network. Ignoring the stochasticity due to the limited size of population, here we focus on these by considering an infinitely large population (i.e., $N \rightarrow \infty$).

Methods

We used Latin hypercube sampling⁵² to identify parameter values which led to model predictions of cycles in incidence which were consistent with those observed. The values of the following parameters were sampled within the ranges listed in Table 1: infectious period (D), duration of immunity (d), degree of contact network (κ), cross-immunity (ψ), effects of primary infection on infectivity (ν) and duration (μ) of a secondary infection. This was done with the function `randomLHS` of the package `lhs` in R computing language⁵⁸.

The basic reproduction number (R_0) was fixed at 1.7 as estimated by⁹. When estimating R_0 , Nguidop-Djomo *et al.*⁹ assumed random mixing among individuals and didn't account for a network structured population. For a network structured population, the random mixing assumption will give rise to an over-estimate of R_0 (Fig. 2 of⁵⁵), so we also consider this effect by assuming that $R_0 = 1.5$ and 1.3 for MP. The total number of infections in our

parameter	Definition	Values or ranges	source
L	Average life span	70*365 days	57
D	Infectious period of a single infection	14–42 days	23
d	Duration of immunity	2–10 years	26
R_0	Basic reproduction number	1.7	9
λ	Transmission rate of single infection	$\lambda = \frac{R_0(d+1)}{D[d(k-2) + (\kappa-1)]}$	55
ψ	Reduction in susceptibility to a secondary infection, resulting from a primary infection (“cross-immunity”)	0.0–1.0	–
ν	Relative infectiousness of a secondary infection, compared to a primary infection.	0.5–2.0	Negative (<1) and positive (>1) effects
μ	Factor by which the duration of a secondary infection differs from that of a primary infection.	0.5–2.0	Negative (<1) and positive (>1) effects
κ	Degree of contact network: number of people with whom one person contacts.	2.5–25.0	assumed

Table 1. Model parameters.

model includes both asymptomatic and symptomatic infections. Since MP may affect all age groups (e.g. ^{6,7,25}), we consider the situation in which the life span is 70 years, the worldwide average life expectancy according to the world fact book⁵⁷.

A value for the interaction parameters ν and μ of 1.0 implies that there is no interaction on the 2nd infection from the primary infection (This is a special situation considered in⁴⁰). If they are less than 1.0, it means the interactions diminish the relevant processes. On the contrary, if they are larger than 1.0, they enhance the processes. The ranges listed in Table 1 allow both increasing and decreasing effects to be selected. To constrain the interactions within biologically reasonable limits, we allow both ν and μ to vary from 0.5 to 2.0. To consider the effect of network-mediated spatial correlation, the contact degree (κ) is allowed to vary from 2.5 to 25.

The Runge-Kutta fourth order method was used to solve the model equations (3–4). As our dynamic system is deterministic, there is one dynamical time series under each set of parameter values. For each time series in which infection persists, weekly rates of new infections with each strain are recorded: the first 800 years were discarded and 200 years were used for analysis. To monitor the time series data and calculate the inter-epidemic periods if periodic changes in the incidence of both strains and total number of infections occur, the spectrum function in the R computing language⁵⁸ is employed. The inter-epidemic period (or the duration of epidemic cycle) will be denoted by EpiT. Following²⁸, the coefficient of variance (CV) of the incidence time series was used to define the shape of epidemic curve and the strength of oscillation in infections over time. It was showed¹⁴ that CV in Japan MP epidemics 1982 to 1990 is about 0.7. In view of this, we regard the epidemics that possess the following characteristics as reasonable approximates to what has been empirically observed in MP epidemics: $0.3 \leq CV \leq 1.4$ and $3 \leq \text{EpiT} \leq 7$ years, along with dominant strain changing among epidemics. In the main text we refer to this as the “characteristically recurrent epidemics of MP”.

References

- Foy, H. M. Infections caused by *Mycoplasma pneumoniae* and possible carrier state in different populations of patients. *Clin. Infect. Dis.* **17**, S37–S46 (1993).
- Korppi, M., Heiskanen-Kosma, T. & Kleemola, M. Incidence of community-acquired pneumonia in children caused by *Mycoplasma pneumoniae*: serological results of a prospective, population-based study in primary healthcare. *Respirology* **9**, 109–114 (2004).
- Dumke, R. *et al.* Optimized serodiagnosis of *Mycoplasma pneumoniae* infections. *Diagn. Microbiol. Infect. Dis.* **73**, 200–203 (2012).
- Waites, K. B. & Talkington, D. F. *Mycoplasma pneumoniae* and its role as a Human pathogen. *Clin. Microbiol. Rev.* **17**, 672–728 (2004).
- Winchell, J. M. *Mycoplasma pneumoniae* – A national public health perspective. *Curr. Pediatr. Rev.* **9**(4), <https://doi.org/10.2174/15733963113099990009> (2013).
- Chalker, V. J. *et al.* Increased incidence of *Mycoplasma pneumoniae* infection in England and Wales in 2010: multilocus variable number tandem repeat analysis typing and macrolide susceptibility. *EuroSurveill.* **16**(19), 19865 (2011).
- Brown, R. J. *et al.* *Mycoplasma pneumoniae* Epidemiology in England and Wales: A national perspective. *Front. Microbiol.* **7**, 157, <https://doi.org/10.3389/fmicb.2016.00157> (2016).
- Jacobs, E. *Mycoplasma pneumoniae*: now in the focus of clinicians and epidemiologists. *EuroSurveill* **17**, 20084 (2012).
- Nguipod-Djomo, P. *et al.* Cyclic epidemics of *Mycoplasma pneumoniae* infections in England and Wales from 1975 to 2009: time-series analysis and mathematical modelling. *Lancet* **382**, S78, [https://doi.org/10.1016/S0140-6736\(13\)62503-9](https://doi.org/10.1016/S0140-6736(13)62503-9) (2013).
- Dorigo-Zetsma, J. W., Dankert, J. & Zaai, S. A. Genotyping of *Mycoplasma pneumoniae* clinical isolates reveals eight P1 subtypes within two genomic groups. *J. Clin. Microbiol.* **38**, 965–970 (2000).
- Pereyre, S. *et al.* The spread of *Mycoplasma pneumoniae* is polyclonal in both an endemic setting in France and in an epidemic setting in Israel. *PLoS ONE* **7**, e38585, <https://doi.org/10.1371/journal.pone.0038585> (2012).
- Dumke, R. & Jacobs, E. Culture-independent multi-locus variable-number tandem-repeat analysis (MLVA) of *Mycoplasma pneumoniae*. *J. Microbiol. Methods* **86**, 393–396 (2011).
- Simmons, W. L. *et al.* Type 1 and type 2 strains of *Mycoplasma pneumoniae* form different biofilms. *Microbiology* **159**, 737–747 (2013).
- Kenri, T. *et al.* Genotyping analysis of *Mycoplasma pneumoniae* clinical strains in Japan between 1995 and 2005: type shift phenomenon of *M. pneumoniae* clinical strains. *J. Med. Microbiol.* **57**, 469–475 (2008).
- Kogoj, R. *et al.* Genetic diversity and macrolide resistance of *Mycoplasma pneumoniae* isolates from two consecutive epidemics in Slovenia. *Eur. J. Clin. Microbiol. Infect. Dis.* **37**, 99–107 (2017).
- Spuesens, E. B. *et al.* Sequence variations in RepMP2/3 and RepMP4 elements reveal intragenomic homologous DNA recombination events in *Mycoplasma pneumoniae*. *Microbiology* **155**, 2182–2196 (2009).

17. Dumke, R., von Baum, H., Lück, P. C. & Jacobs, E. Subtypes and variants of *Mycoplasma pneumoniae*: local and temporal changes in Germany 2003–2006 and absence of a correlation between the genotype in the respiratory tract and the occurrence of genotype-specific antibodies in the sera of infected patients. *Epidemiol. Infect.* **138**, 1829–1837 (2010).
18. Martinez, M. A. *et al.* Identification of P1 types and variants of *Mycoplasma pneumoniae* during an epidemics in Chile. *J. Med. Microbiol.* **59**, 925–929 (2010).
19. Zhao, F. *et al.* Culture-independent detection and genotyping of *Mycoplasma pneumoniae* in clinical specimens from Beijing China. *PLoS ONE* **10**, e0141702, <https://doi.org/10.1371/journal.pone.0141702> (2015).
20. Brown, R. J., Holden, M. T., Spiller, O. B. & Chalker, V. J. Development of a multilocus sequence typing scheme for molecular typing of *Mycoplasma pneumoniae*. *J. Clin. Microbiol.* **53**, 3195–3203 (2015).
21. Dumke, R. & Jacobs, E. Antibody Response to *Mycoplasma pneumoniae*: Protection of host and influence on outbreaks? *Front. Microbiol.* **7**, 39 (2016).
22. Anderson, R. M. & May, R. M. *Infectious Disease of Humans: Dynamics and Control*. (Oxford University Press, Oxford, 1991).
23. Clyde, W. A. Clinical overview of typical *Mycoplasma-Pneumoniae* infections. *Clin. Infect. Dis.* **17**, S32–S36 (1993).
24. Foy, H. M. *et al.* Second attacks of pneumonia due to *Mycoplasma pneumoniae*. *J. Infect. Dis.* **135**, 673–677 (1977).
25. Ito, I. *et al.* Culturally verified *Mycoplasma pneumoniae* pneumonia in Japan: a long-term observation from 1979–99. *Epidemiol. Infect.* **127**, 365–367 (2001).
26. Lind, K. *et al.* A seroepidemiological study of *Mycoplasma pneumoniae* infections in Denmark over 50-year period 1946–1995. *Eur. J. Epidemiol.* **13**, 581–586 (1997).
27. Keeling, M. J. & Rohani, P. *Modelling Infectious Disease in Humans and Animals*. Princeton, NJ. (Princeton University Press, USA, 2008).
28. Omori, R. *et al.* The determinant of periodicity in *Mycoplasma pneumoniae* incidence: an insight from mathematical modelling. *Sci. Rep.* **5**, 14473, <https://doi.org/10.1038/SREP14473> (2015).
29. Nakata, Y. & Omori, R. Delay equation formulation for an epidemic model with waning immunity: an application to *Mycoplasma pneumoniae*. *IFAC-PapersOnLine* **48**, 132–135 (2015).
30. Woodhead, M. & Macfarlane, J. Local antibiotic guidelines for adult community-acquired pneumonia (CAP): a survey of UK hospital practice in 1999. *J. Antimicrob. Chemother.* **46**, 141–143 (2000).
31. Morozumi, M. *et al.* Increased macrolide resistance of *Mycoplasma pneumoniae* in paediatric patients with community acquired pneumonia. *Antimicrob Agents Chemother* **52**, 348–350 (2008).
32. Zhao, F. *et al.* Surveillance of macrolide-resistant *Mycoplasma pneumoniae* in Beijing, China, from 2008 to 2012. *Antimicrob. Agents Chemother.* **57**, 1521–1523 (2013).
33. Suzuki, Y. *et al.* Gene mutations associated with macrolide-resistance and p1 genotyping of *Mycoplasma pneumoniae* isolated in Yamagata, Japan, between 2004 and 2013. *Kansenshogaku Zasshi* **89**, 16–22 (2015).
34. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74, <https://doi.org/10.1371/journal.pmed.0050074> (2008).
35. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys Rev Lett* **86**, 3200–3, <https://doi.org/10.1103/PhysRevLett.86.3200> (2001).
36. Watts, D. J. & Strogatz, S. H. Collective dynamics of small world networks. *Nature* **393**, 440–442 (1998).
37. Leventhal, G. E., Hill, A. L., Nowak, M. A. & Bonhoeffer, S. Evolution and emergence of infectious diseases in theoretical and real-world network. *Nat. Commun.* **6**, 6101, <https://doi.org/10.1038/ncomms7101> (2015).
38. Diekmann, O., Heesterbeek, H. & Britton, T. *Mathematical Tools for Understanding Infectious Disease Dynamics*. (Princeton University Press, Princeton and Oxford, 2013).
39. Rozhnova, G. & Nunes, A. Fluctuation and oscillations in a simple epidemic model. *Phys Rev E* **79**, 041922 (2009).
40. Zhang, X.-S. Epidemic cycling in a multi-strain SIRS epidemic network model. *Theor. Biol. Med. Model.* **13**, 14 (2016).
41. Zhang, X.-S. & Cao, K.-F. The impact of coinfections and their simultaneous transmission on antigenic diversity and epidemic cycling of infectious diseases. *BioMed Res. Int.* **375862**, <https://doi.org/10.1155/2014/375862> (2014).
42. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013).
43. Danon, L., Read, J. M., House, T. A., Vernon, M. C. & Keeling, M. J. Social encounter networks: characterizing Great Britain. *Proc R Soc B* **280**, 20131037, <https://doi.org/10.1098/rspb.2013.1037> (2013).
44. Kwok, K. O., Cowling, B., Wei, V., Riley, S. & Read, J. M. Temporal variation of human encounters and the number of locations in which they occur: a longitudinal study of Hong Kong residents. *J. R. Soc. Interface* **15**, 20170838, <https://doi.org/10.1098/rsif.2017.0838> (2018).
45. Ferguson, N., Anderson, R. & Gupta, S. The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proc Natl Acad Sci USA* **96**, 790–794 (1999).
46. Halstead, S. B., Mahalingam, S., Marovich, M. A., Ubol, S. & Mosser, D. M. Intrinsic antibody-dependent enhancement of microbial infection in macrophages: disease regulation by immune complexes. *Lancet Infect Dis.* **10**, 712–722, [https://doi.org/10.1016/S1473-3099\(10\)70166-3](https://doi.org/10.1016/S1473-3099(10)70166-3) (2010).
47. Song, Q., Xu, B.-P. & Shen, K.-L. Effects of bacterial and viral co-infections of *Mycoplasma pneumoniae* in children: analysis report from Beijing Children's hospital between 2010 and 2014. *Int. J. Clin. Exp. Med.* **8**, 15666–15674 (2015).
48. Tsiodras, S. *et al.* Central nervous system manifestations of *Mycoplasma pneumoniae* infections. *J. Infect.* **51**, 343–354 (2005).
49. Goncalves, S., Abramson, G. & Gomes, M. F. C. Oscillations in SIRS model with distributed delays. *Eur. Phys. J. B* **81**, 363–371 (2011).
50. Bansal, S., Grenfell, B. & Meyers, L. A. When individual behaviours matters: homogeneous and network models in epidemiology. *J.R. Soc. Interface* **4**, 879–891 (2007).
51. Guclu, H. *et al.* Social Contact Networks and Mixing among Students in K-12 Schools in Pittsburgh, PA. *PLOS One* **11**(3), e0151139, <https://doi.org/10.1371/journal.pone.0151139> (2016).
52. Iman, R. L., Helton, J. C. & Campbell, J. E. An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment. *J. Qual. Technol.* **13**, 174–183 (1981).
53. Saltelli, A. *et al.* Variance based sensitivity analysis of model output. *Design and estimator for the total sensitivity index. Comput Phys Commun.* **181**, 259–270 (2010).
54. Hughes, J. *et al.* Benefits and unintended consequences of antimicrobial de-escalation: Implications for stewardship programs. *PLOS ONE* **12**, e0171218 (2017).
55. Eames, K. T. D. & Keeling, M. J. Modelling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc Natl Acad Sci USA* **99**, 13330–13335 (2002).
56. Fan, L. *et al.* Pediatric clinical features of *Mycoplasma pneumoniae* infection are associated with bacterial P1 genotype. *Exp Ther Med* **14**, 1892–1898 (2017).
57. The World Fact-book Life Expectancy. Cia.gov. <https://www.cia.gov/library/publications/the-world-factbook/fields/2102.html>, (2012).
58. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/> (2016).

Acknowledgements

The work was carried out whilst the authors were employed by Public Health England. We thank Elaine Stanford for her kind support for this work. X.-S.Z. would like to dedicate the paper to Professor Shou-Li Peng of Yunnan University who passed away on 18/08/2018 for his teachings, support and encouragements.

Author Contributions

Conceived and designed the mathematical model and analysis: X.-S.Z., H.Z., E.V., V.C.; performed the study: X.-S.Z.; wrote the paper: X.-S.Z., H.Z., E.V., V.C.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36325-z>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019