

Electronic Supplementary Material

Juan Pablo Aparicio and Mercedes Pascual

1 Basic Reproduction Numbers

We define the basic reproduction number as the expected number of cases produced by an average infected individual in the second generation. In all cases we assume that the probability of transmission per contact in an interval δt is $1 - \exp(-\tau\delta t)$, and that infected individuals are infectious with infectious period exponentially distributed with mean $1/\gamma$. The removal rate $\gamma = r + \mu$ has two contributions: the recovery rate r and the mortality rate μ . All the rates are measured in units of the inverse of the recovery rate and therefore are non-dimensional with $r^{-1} = 1$.

1.1 The basic reproduction number for homogeneously mixed networks

The probability that an infectious individual infects a susceptible contact during the infectious period is $\rho = \frac{\tau}{\tau + \gamma}$. One average infectious individual placed in an homogeneously mixed population of size $N - 1$ will produce $(N - 1)\rho$ secondary infections. Secondary cases compete among themselves for the remaining susceptible population. By disregarding such competition and re-infections, we estimate that a secondary case produces $(N - 1)(1 - \rho)\rho$ new cases. Because R_0 is finite, ρ goes to zero as N increases, in such a way that $(N - 1)(1 - \rho)\rho$ remains constant (and therefore $\tau \rightarrow 0$ as $N \rightarrow \infty$). For large N , the basic reproduction number is therefore given by

$$R_{0HM} = N\rho \simeq N\frac{\tau}{\gamma}. \quad (1)$$

By defining the population-level transmission parameter as $\beta \equiv N\tau$, we can write the mean field model as the mass-action equation $dI/dt = (\tau S - \gamma)I = \beta SI/N - \gamma I$ with threshold β/γ equal to the basic reproduction number (1).

1.2 The basic reproduction number for Poisson random networks

A secondary case has i contacts with probability $\Pi(i) = iP(i)/n$, where $P(i)$ is the degree distribution and n , its mean. The probability of transmission per contact is $\rho \equiv \frac{\tau}{\tau+\gamma}$ and therefore the basic reproduction number becomes (Andersson 1997, Diekmann & Heesterbeek 2000)

$$R_{0rdm} = \sum_{i=1}^n (i-1)\Pi(i)\rho = \left[(n-1) + \frac{VAR(i)}{n} \right] \rho,$$

where we assumed that the source case would not become a susceptible contact of the secondary cases. For a Poisson network, $VAR(i) = n$ and therefore

$$R_{0rdm} = n\rho. \quad (2)$$

1.3 Basic reproduction number for regular networks

For $\phi = 0$ each individual is in contact with eight neighbors. These contacts are of two types, four in the corners (hereafter denoted as type A), and the remaining four in between these (denoted as type B). For simplicity, we assume that secondary cases compete among themselves for susceptible contacts but do not compete with other generations of infecteds (such as the source case itself or the ternary cases). The source case produces i secondary cases with probability $\rho^i(1-\rho)^{n-i}$. For each of these cases, several different configurations are possible and therefore, we can compute the approximate expected number of ternary cases produced by the secondary cases which results in a polynomial of degree 11. Here we present a much simpler approximation which works equally well.

C			
D	A	B'	A''
E	B	0	
	A'		

Figure 1: Diagram of the different types of secondary cases to illustrate competition for susceptibles. A secondary case of type A competes with secondary cases of type B or C because their respective neighborhoods overlap.

Instead of computing the probabilities corresponding to the different configurations, we consider that a fraction ρ of the contacts of the source are infected. Then, if we have a secondary case of type A (fig 1), we assume that A is infected with ‘level’ 1, while node B is infected with ‘level’ ρ . Then node D is infected either by A or B in proportions $\frac{1}{1+\rho}$ and $\frac{\rho}{1+\rho}$ respectively. Thus, the average number of ternary cases of type D produced by secondary cases of type A is given approximately by $\frac{\rho}{1+\rho}$.

More specifically, the contribution to R_0 from a secondary case of type A can be estimated as follows:

A secondary case A will infect contact C with probability ρ (since there is no competition with other secondary cases).

The average number of ternary cases of type D produced by secondary cases of type A is given by $\frac{\rho}{1+\rho}$.

A secondary case of type A will infect a contact E with probability $\frac{\rho}{1+2\rho}$ because E is shared with both B and A' .

Finally, the secondary case A will infect contacts B or B' with probability $\rho(1 - \rho)$.

From all these contributions, the expected number of ternary cases produced by a secondary case of type A is:

$$R_{0A} = \rho + \frac{2\rho}{1+\rho} + \frac{2\rho}{1+2\rho} + 2(1-\rho)\rho = \rho \left(3 - 2\rho + \frac{2}{1+\rho} + \frac{2}{1+2\rho} \right)$$

In a similar fashion, the expected number of ternary cases produced by a secondary case of type B is:

$$R_{0B} = \frac{2\rho}{1+\rho} + \frac{\rho}{1+2\rho} + 4(1-\rho)\rho = \rho \left(4 - 4\rho + \frac{2}{1+\rho} + \frac{1}{1+2\rho} \right).$$

Finally, the basic reproduction number becomes

$$R_{0sp} = \frac{1}{2}(R_{0A} + R_{0B}) = \frac{1}{2}\rho \left(7 - 6\rho + \frac{4}{1+\rho} + \frac{3}{1+2\rho} \right) \quad (3)$$

In Fig. 2, we compare expression 3 with empirical values obtained for different values of τ (see also section 1.5 below).

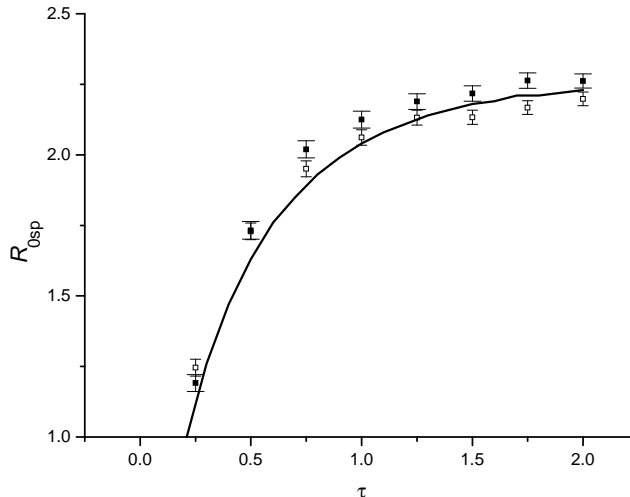


Figure 2: Empirical estimates of R_{0sp} obtained with 5000 simulations for $\mu = 0$ (open squares) and $\mu = 0.05$ (solid squares). Bars represent approximate 95% confidence intervals. The continuous line corresponds to equation 3 with $\mu = 0$.

1.4 The basic reproduction number for small-world networks

An average individual in the small-world network will have n contacts. The probability that such individual has i long distance contacts depends on the disorder parameter ϕ and is approximately binomial,

$$P(i, \phi) = \binom{n}{i} (1 - \phi)^{n-i} \phi^i.$$

In other words, an average individual will have i long-distance contacts (and $n - i$ local contacts) with probability $P(i, \phi)$. Long distance cases will not compete in the limit of an infinite size network for susceptible individuals with other cases. We will assume that a long distance contact produces R_{0rdm} infections among its contacts¹.

A fact that complicates the computation of R_0 for small-world networks is that the number of ternary cases produced by a secondary case depends on the type and number of contacts of the index case itself. For example, suppose that all of the contacts of the index case are local contacts (the index

¹For small ϕ there is little variability in the distribution of the number of contacts of the nodes, and therefore $R_{0rdm} \sim (n - 1)\rho$, while for a Poisson random network $R_{0rdm} = n\rho$. Here we used $R_{0rdm} = n\rho$.

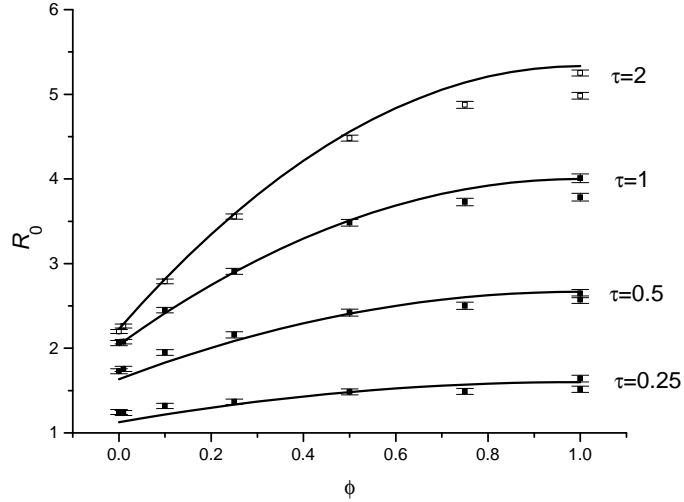


Figure 3: Comparison of empirical estimates and analytical calculations of R_{0sw} for different values of the disorder parameter ϕ . A total of 5000 stochastic simulations were performed for $\mu = 0$ in small-world networks, with 50 different network configurations and 100 realizations of the model for each configuration. Bars represent approximate 95% confidence intervals. The continuous line corresponds to equation 4 with $\mu = 0$. Two empirical estimates are shown for $\phi = 1$: the lower value was obtained with small world networks constructed with this value of the disorder parameter, the top value was obtained with Poisson random networks. Both cases are shown because, for this value of ϕ , the networks constructed with the small world algorithm are not exactly Poisson random networks. The algorithm to generate small world networks is a two-dimensional version of the one described by Watts and Strogatz (1998, see also Roy & Pascual 2005), in which the starting point is a spatial grid with neighborhoods composed of the eight near neighbors and periodic boundary conditions. Each local connection is rewired with probability ϕ to a random site, avoiding self and multiple connections. Only those configurations that are completely connected are kept.

case does not have long distance contacts). The secondary case itself may have some long distance contacts (in this case, no more than five, but for simplicity,² we assume that the maximum number is n). With probability $P(0)$, a secondary case will have only local contacts. In this case it will produce, on average R_{0sp} ternary cases (this is equivalent to the case in which all of the contacts of the index case have only contacts of the local type). With probability $P(2)$, a secondary case will have two long distance contacts (or only two of the contacts of the index case will have one long distance contact). In general, a secondary case will have i long distance contacts with probability $P(i)$.

Then, if the index case has only local contacts (which occurs with probability $P(0)$), but there are in turn i long distance contacts among the contacts of these local neighbors, the expected number of ternary cases is no longer equal to R_{0sp} . This number may instead be estimated as follows. Given that a secondary case becomes infected, it will either have only local contacts with probability $(n - i)/n$, or, a long distance contact with probability i/n . It will produce, respectively for these two different cases, an average of R_{0sp} and R_{0rdm} ternary infections; By averaging over all values of i , we obtain

$$R_0(0) = \sum_{i=0}^n P(i, \phi) \left[\frac{(n - i)}{n} R_{0sp} + \frac{i}{n} R_{0rdm} \right]$$

when the index case has only local contacts.

Consider now the case where the index infection has only one long distance contact, which occurs with probability $P(1)$. The index case will infect its long distance contact with probability proportional to $1/n$. But again, the $n - 1$ local contacts of the index case may have themselves long distance contacts. As in the previous calculation, if there are i long distance contacts among the local contacts of the index case, then the expected number of ternary cases produced by these is given by

$$R_{0av} = \sum_{i=0}^n P(i, \phi) [(n - i)R_{0sp} + iR_{0rdm}] / n.$$

However, if the secondary case is the long distance contact of the index case, it will produce R_{0rdm} ternary cases. Therefore, the total number of

²For small ϕ , $P(i > 5) \simeq 0$, while for large ϕ , $P(0) \simeq 0$. Therefore, n is a fairly good approximation.

ternary cases produced by secondary cases of an index infection with only one long distance contact becomes:

$$R_0(1) = [(n - 1)R_{0av} + R_{0rdm}]/n.$$

In general, if the index case have exactly i long distance contacts, the number of ternary cases produced by an average secondary case is given by

$$R_0(i) = [(n - i)R_{0av} + iR_{0rdm}]/n.$$

The basic reproduction number for the small-world network is given by the corresponding average over all of these possibilities

$$R_{0sw}(\phi) = \sum_{i=0}^n P(i, \phi) R_0(i) = \sum_{i=0}^n P(i, \phi) [(n - i)R_{0av} + iR_{0rdm}]/n \quad (4)$$

Values obtained with equation (4) are in excellent agreement with empirical estimations of the basic reproduction number obtained from simulations (see Fig. 3).

1.5 Empirical estimates of the basic reproductive number

To evaluate the above expressions for the basic reproduction numbers, we obtained estimates from network simulations as follows. For a given value of ϕ , we created a network configuration, introduced one infected individual in a random site, and calculated a value of R_0 as the number of ternary cases over the number of secondary cases. For each value of ϕ (0, 0.01, 0.1, 0.25, 0.5, 0.75, 1), the process was repeated for 50 different network configurations and 100 stochastic simulations each. The value of R_0 was estimated as the mean of the 5000 simulations. An excellent agreement between empirical estimates and expression (4) was observed (see Fig. 3).

It should be noticed that as ϕ reaches one, expression (4) underestimates the empirical value of R_0 , which follows from the fact that the standard algorithm used to build the small-world network does not produce a Poisson network for $\phi = 1$. Thus, in Fig. 3 we also include the empirical estimate of R_0 obtained with Poisson random networks.

2 Empirical Parameterization

At the beginning of the epidemic, new infections are produced at rate $n_{eff}\tau I(t)$ per unit of time, with $I(t) = I_0 e^{\lambda t}$. Therefore the incidence at a given time $t \geq D$ is given by

$$Inc(t) = \int_{t-D}^t n_{eff}\tau I(s) ds = \frac{n_{eff}\tau I_0}{\lambda} (e^{-\lambda D} - 1) e^{\lambda t} \quad (5)$$

where I_0 is the (initial) infected population at $t = 0$, and D is the period during which the new cases are cumulated for reporting. From the linear regression of the data for $\ln Inc(t)$ vs t , estimates are obtained of the slope ($\hat{\lambda}$) and y-intercept ($b \equiv \ln Inc_0$, where Inc_0 is back-extrapolated).

The best fit of model solutions $I(t) = I_0 e^{\lambda t}$ to the data are obtained when $\lambda = \hat{\lambda}$ and

$$I_0 = e^{\hat{\lambda} D} Inc_0 \left(\frac{\hat{\lambda}}{n_{eff}\tau} \right) \left(\frac{1}{e^{\hat{\lambda} D} - 1} \right). \quad (6)$$

For the standard mean field model, $\lambda = n_{eff}\tau - \gamma$ and therefore

$$n_{eff}\tau \simeq \hat{\lambda} + \hat{\gamma}$$

where $\hat{\gamma}$ is an estimate of γ .

For the modified mean field model, $\lambda = (\tau + \gamma)(R_0 - 1)$ and then

$$R_0 \simeq \frac{\hat{\lambda}}{\hat{\tau} + \hat{\gamma}} + 1$$

with $\hat{\tau}$ an estimate of τ .

For both cases, the effects of network structure on the initial spread of the disease are taken into account implicitly in n_{eff} and R_0 respectively. These values may be estimated from data when estimates of individual level parameters (specifically γ and τ) are available.

3 Comparison of the stochastic simulations with ODE models

Here we compare the stochastic simulations with the trajectories of the ODEs for both the standard mean-field (MF, eq. 1 in main text) and the modified

mean-field (MMF, eqs. 2-5 in main text) model as we vary the transmission rate τ . Six different quantities are compared that characterize the long-term and initial transients of the system: the equilibrium of susceptibles (left top panel Fig. 4), the equilibrium of infecteds (right top panel), the value of the first maximum of infecteds or peak of the first epidemic (left middle panel), the timing of this peak (right middle panel), the value of the first minimum of infecteds or the trough of the first epidemic (left bottom panel), and the timing of this trough (right bottom panels). A total of 300 stochastic runs were considered for each value of τ , and logarithms of susceptibles and infecteds numbers were used. For each run a difference was computed between the stochastic quantity of interest (indicated as the label of the y axis) and the same quantity in the corresponding ODE simulation. The plots show the average of these differences for the MF (eq. 1 in main text) and for the MMF system (eqs. 2-5 in main text), as well as the standard deviation around these means when sufficiently large to be seen in the graph. Negative differences indicate that the ODE overestimates the corresponding quantity in the stochastic system, and vice versa. For all quantities and most values of τ , the MMF model provides a better approximation to the stochastic run than the standard MF model. The only exception is for the value of the first trough for large values of $\tau \geq 1.75$. However, even in this case, the timing of the trough is better approximated by the MMF which also captures better the overall oscillatory pattern of the transients (see Fig. 5).

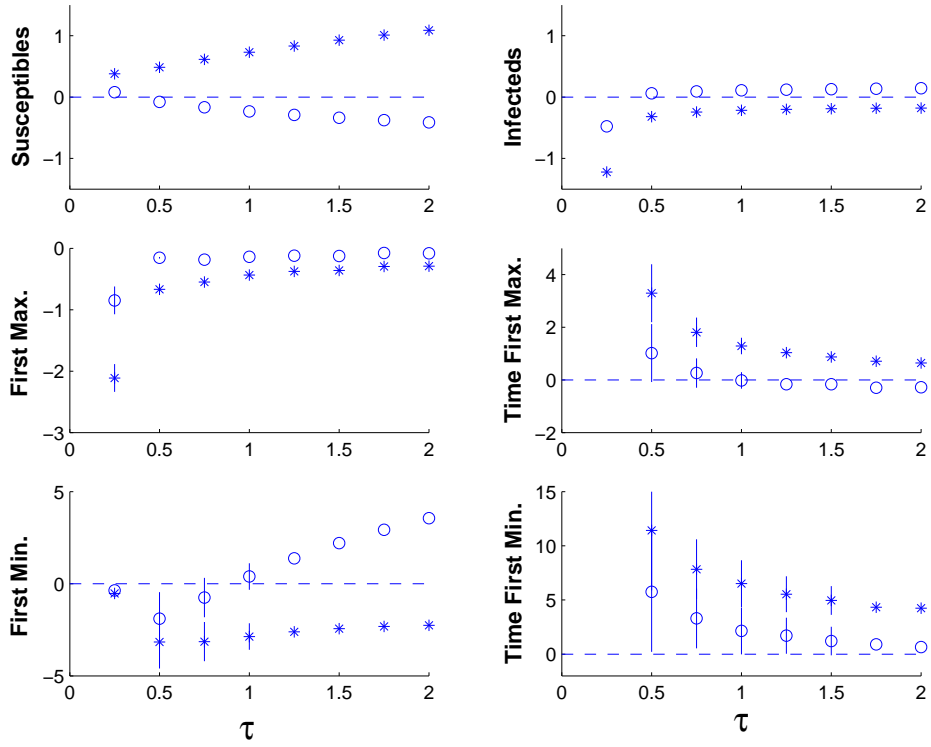


Figure 4: These panels show the comparison of the stochastic simulations with the trajectories of the ODEs for both the standard mean-field (eq. 1 in main text, stars) and the modified mean-field model (eqs. 2-5 in main text, circles) for different values of the transmission rate τ . See section "Comparison of the stochastic simulations with ODE models" for details. In all simulations, $r = 1$, $\mu = 0.05$, and $N = 90000$.

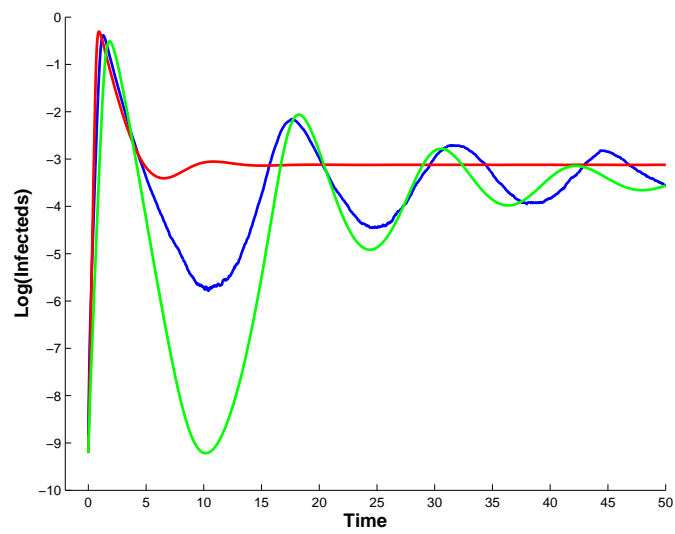


Figure 5: Logarithm of infected numbers for the stochastic model (in blue), the standard mean field model (eq. 1 in main text, in red) and the modified mean field model (eqs. 2-5 in main text, in green) for $\tau = 2$. Other parameters as in Fig. 4.