The Impact of Contact Tracing in Clustered Populations

Thomas House*, Matt J. Keeling

Warwick Mathematics Institute and Department of Biological Sciences, University of Warwick, Coventry, United Kingdom

Abstract

The tracing of potentially infectious contacts has become an important part of the control strategy for many infectious diseases, from early cases of novel infections to endemic sexually transmitted infections. Here, we make use of mathematical models to consider the case of partner notification for sexually transmitted infection, however these models are sufficiently simple to allow more general conclusions to be drawn. We show that, when contact network structure is considered in addition to contact tracing, standard "mass action" models are generally inadequate. To consider the impact of mutual contacts (specifically clustering) we develop an improvement to existing pairwise network models, which we use to demonstrate that *ceteris paribus*, clustering improves the efficacy of contact tracing for a large region of parameter space. This result is sometimes reversed, however, for the case of highly effective contact tracing. We also develop stochastic simulations for comparison, using simple re-wiring methods that allow the generation of appropriate comparator networks. In this way we contribute to the general theory of network-based interventions against infectious disease.

Introduction

Modelling has become a central tool in understanding the epidemiology of infectious disease, and designing control strategies. One control method, contact tracing, has been considered in a large number of disease contexts. These include the 2003 SARS pandemic [1,2], the 2001 UK FMD epidemic [3–6], contingency planning for deliberate release of smallpox [7,8], and control of sexually transmitted infections [9–11]. A particular benefit of tracing is that it allows targeting of control, at the cost of effort spent on finding the individuals at risk.

Since contact tracing takes place as a process over the network of interactions between hosts, it is natural to consider network-based models of this process. Theoretical work has so far dealt with contact tracing as a branching process [12], through modifications to mean-field equations [13], pairwise approximations [14] and simulation [15]. This work means that the implications of heterogeneous numbers of contacts (and related network properties such as assortativity) for the efficacy of contact tracing are reasonably well understood.

For the case of clustering, due to the analytical challenge posed by the existence of short closed loops in the contact network, it has generally been more difficult to make similar progress. Existing theoretical work has therefore either been restricted to the 'limiting case' of clump structured populations, with all clustering due to completely connected cliques [16], or else simulation on exemplar networks [13,14,17].

In this work, we derive an improved triple closure for clustered pairwise models that removes two significant problems with existing closure regimes, and use this to make a systematic investigation of the impact of clustering on the efficacy of contact tracing, keeping other network and epidemiological parameters constant as appropriate. We find that, for many parameter choices, there are intuitive explanations, borne out by modelling, for the increased impact of contact tracing as clustering increases. This is not, however, a completely general result, meaning that the full implications of clustering for the efficacy of contact tracing are subtle and should be the subject of case by case investigation.

We perform our analysis within the *SIS* paradigm, meaning that while some of our terminology will be general to all infectious disease epidemiology, other statements will be geared towards the modelling of sexually transmitted infections where recovery/treatment does not confer lasting immunity.

Methods

Modelling contact tracing

The dynamics underpinning our model are shown schematically in Figure 1. Individuals are either susceptible (*S*), infectious (*I*) or traced (*T*) and move between these compartments due to four processes: infection; treatment; tracing; and stopping tracing. This paradigm is suitable for the consideration of sexually transmitted diseases, where infectious individuals can transmit infection to contacts, then seek treatment, which clears the pathogen and stops transmission but leaves the individual susceptible. It also involves the process of contact tracing, which we use as a general term that includes both partner notification and efforts by public-health workers to track down potentially infected individuals.

The four processes described so far separate into two categories: those that happen at an individual level, and contact processes. Seeking treatment and the cessation of tracing take place in the population at rates proportional to a number of individuals, and so fall into the former category. Using square brackets around a quantity to indicate its expected number in the population (so that quantities in square brackets are extensive expected numbers rather than intensive proportions) we take treatment to happen at a rate $g[I]$, where g is the

Author Summary

There are multiple ways to control infectious diseases—vaccination and drugs such as antibiotics or anti-virals form part of the pharmaceutical approach, however another route is to stop people infecting each other. This can be done either through general efforts to reduce epidemiologically relevant contacts, or through a more targeted attempt to trace the contacts of known cases who can then be isolated or treated. The impact of this kind of contact tracing is *a priori* likely to depend strongly on the network of contacts linking people together. In this paper, we develop new mathematical and computational techniques to model the impact of clustering: the probability that any two contacts of a given individual are also linked to each other in the network, creating triangles. Often, and for intuitively understandable reasons, the presence of clustering increases the efficacy of contact tracing, however we show that in the regime of highly effective contact tracing sometimes the opposite is true.

treatment rate constant, and the cessation of efforts to trace a treated individual's contacts to happen at rate $g_T[T]$, where g_T is the rate constant associated with the end of tracing.

Infection and contact tracing, on the other hand, are contact processes, and so take place at a rate proportional to a number of partnerships in the population. The full set of partnership links can be thought of as forming a network, through which contact processes spread. For infection, the rate is $\tau[S \leftarrow I]$, where the term in brackets is the number of susceptible-infectious pairs in the population and τ is the transmission rate constant, and for tracing, the rate is $\rho[I \leftarrow T]$, where the term in brackets is the number of infectious-traced pairs in the population and ρ is the tracing rate constant. We have introduced here a notation in which a arrow is drawn from a state that transmits across the link to the state that is

affected by the transmission, which will become important when we consider triples in addition to pairs.

To consider the impact of network structure, in particular clustering, on the efficacy of contact tracing, we consider a scenario in which an infection with underlying *SIS* dynamics is at its endemic equilibrium, and then contact tracing is introduced and the numbers infectious measured over time. This requires a dynamical model, and so we now turn to two complementary methods that we use to study the system in question: ODE-based models and stochastic simulation.

ODE-based models

Models based on ordinary differential equations (ODEs) are widely used in infectious disease modelling. We present here a series of ODE systems that can be used in the context of network models, starting with mean-field approaches, and moving on to pairwise models. We have found that, for application to contact tracing, mean-field models and existing pairwise closures are inadequate and so we develop an improved pairwise model to study this system.

Mean-field models. For *SIS* dynamics with transmission rate τ across a network link and treatment rate g on a large network, the expected numbers of susceptible and infectious individuals evolve according to the following exact, but unclosed, set of equations.

$$\begin{aligned}
\frac{d}{dt}[S] &= -\tau[S \leftarrow I] + g[I], \\
\frac{d}{dt}[I] &= \tau[S \leftarrow I] - g[I].
\end{aligned} \tag{1}$$

In our notation, $[A]$ refers to the number of individuals in state A, $[A-B]$ and $[A \leftarrow B]$ to the number of pairs with one individual in state A and one in state B, and a directed arrow on the right hand side of a differential equation denotes the direction of transmission for a contact process.
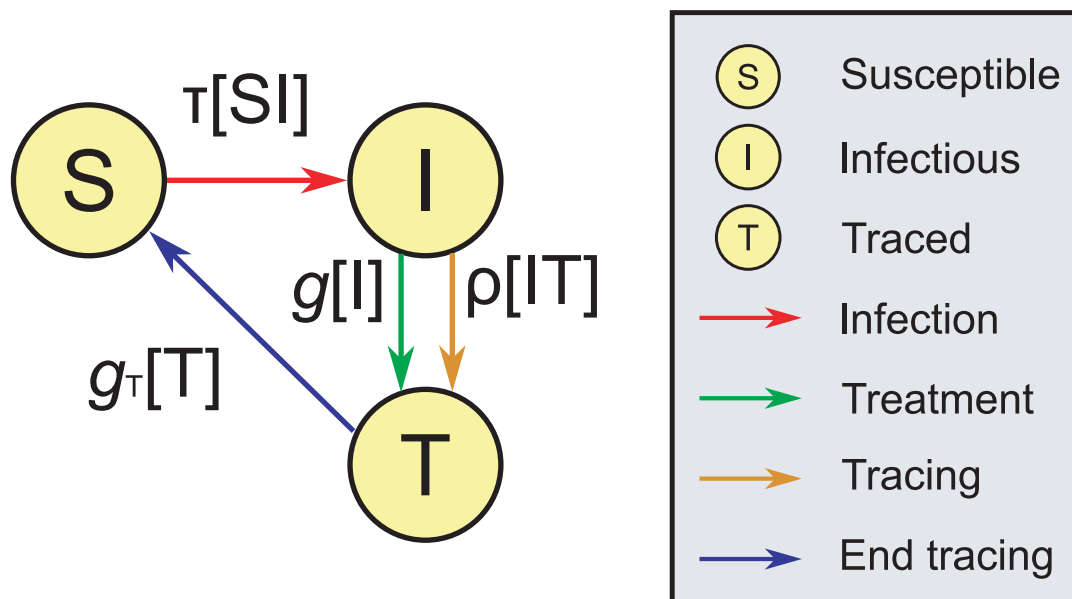

Figure 1. The compartments and processes for *SIS* dynamics with contact tracing. Processes are shown with coloured arrows, labelled according to the rate at which they happen.
doi:10.1371/journal.pcbi.1000721.g001

To produce a mean-field model, we use a low-level closure that approximates pairs in terms of individuals.

$$[A \leftarrow B] \approx \frac{n}{N}[A][B], \qquad (2)$$

where N is the number of nodes in the network, and n is the number of links per node. For *SIR* dynamics, improvements of this scheme are possible that have a factor of $n-1$ in the numerator of (2) in the place of n, representing the fact that after the first infection, each infected individual in an unclustered network will have one fewer link due to the individual they were infected by. For clustered networks, *SIS* dynamics and contact tracing, all of which we are considering here, it is not clear that a similar argument can be used and so we keep the factor of n.

Pairwise models. In pairwise models, rather than using assumptions like (2), equations for the pair-level variables that appear on the right-hand side of (1) are written down, leading to triple-level variables that are then closed in terms of pairs.

The starting point for our analysis is the standard pairwise model for *SIS* dynamics [18], with transmission rate τ across a network link and treatment rate g. This consists of the unclosed equations (1) above, together with the following equations for pairs.

$$\frac{d}{dt}[S-S] = -2\tau[S-S \leftarrow I] + 2g[S-I],$$
$$\frac{d}{dt}[S-I] = \tau([S-S \leftarrow I] - [I-S \leftarrow I] - [S \leftarrow I]) + g([I-I] - [S-I]), \quad (3)$$
$$\frac{d}{dt}[I-I] = 2\tau([I-S \leftarrow I] + [S \leftarrow I]) - 2g[I-I].$$

Here, $[A-B \leftarrow C]$ is the number of nodes of type B connected to both an A and a C, which may or may not be connected themselves. We have continued use of the notation in which a directed arrow on the right hand side of a differential equation denotes the direction of transmission for a contact process, as explained above.

The equations (3) are, like (1), exact, but to produce an integrable system it is necessary to introduce a system of spatial closure. The standard approximation for a population of size N, with exactly n links per node and a clustering coefficient of ϕ—defined as the ratio of triangles to triples in the network—was derived in [19] and is:

$$[A-B \leftarrow C]$$
$$\approx \frac{n-1}{n}\left((1-\phi)\frac{[A-B][B-C]}{[B]} + \phi\frac{N}{n}\frac{[A-B][B-C][C-A]}{[A][B][C]}\right). \quad (4)$$

For clarity about the definition of ϕ, where the network adjacency matrix is $\mathbf{G} = (G_{ij})$, then

$$\phi = \frac{\mathrm{Trace}(\mathbf{G}^3)}{\left(\sum_{i,j}(\mathbf{G}^2)_{ij}\right) - \mathrm{Trace}(\mathbf{G}^2)} \in [0,1]. \quad (5)$$

There are two problems with the approximation (4) that are particularly relevant for the question of contact tracing. The first of these is that we would like to preserve the pair-level relation $\sum_b[A-b] = n[A]$. For the dynamical system (3), this pair-level relation will be preserved over time provided the following triple-level equation holds:

$$\sum_a[a-B \leftarrow C] = (n-1)[B-C]. \quad (6)$$

Equation (6) above holds for the standard closure for unclustered networks, but fails to be satisfied for non-zero clustering. The second problem with the standard closure is the question of how triangles of three infected individuals behave during the early asymptotic stage of an epidemic, where all dynamical variables are governed by the proportion of the population that is infectious, $([I]/N) \ll 1$. While for pure *SIS* dynamics these triples are not dynamically important, when we come to consider contact tracing similar terms will become relevant. Under (4) and assuming the prevalence of infection is very low, the proportion of unclosed triples composed of three infected individuals is proportional to $[I]/N$ as expected. However, under (4), the proportion of triangles where all three individuals are infected is not small (and does not scale with $[I]/N$); clearly, this is inconsistent and should be rectified in any improved closure.

Motivated by these two considerations, we propose an alternative that respects (6) and has appropriate polynomial dependence on $[I]/N$ during the early epidemic.

$$[A-B \leftarrow C]$$
$$\approx (n-1)\left((1-\phi)\frac{1}{n}\frac{[A-B][B-C]}{[B]} + \phi\frac{[A-B][B-C][C-A]}{[A]\sum_a([a-B][a-C]/[a])}\right). \quad (7)$$

This closure breaks the standard symmetry between A and C, however if contact processes are consistently identified on the right hand side of ODE systems like (3) using arrows, then this is not conceptually problematic.

For the rest of this paper, we call pairwise models based on the closure (7) improved pairwise models; while pairwise models based on (4) are called standard pairwise models.

Full pairwise system. Putting together all four processes for our model with tracing, our pairwise system consists of the following exact equations together with the closure approximation (7).

$$\frac{d}{dt}[S] = -\tau[S \leftarrow I] + g_T[T],$$
$$\frac{d}{dt}[I] = \tau[S \leftarrow I] - \rho[I \leftarrow T] - g[I],$$
$$\frac{d}{dt}[T] = \rho[I \leftarrow T] + g[I] - g_T[T],$$
$$\frac{d}{dt}[S-S] = -2\tau[S-S \leftarrow I] + 2g_T[S-T],$$
$$\frac{d}{dt}[S-I] = \tau([S-S \leftarrow I] - [I-S \leftarrow I] - [S \leftarrow I])$$
$$\qquad - \rho[S-I \leftarrow T] - g[S-I] + g_T[I-T],$$
$$\frac{d}{dt}[I-I] = 2\tau([I-S \leftarrow I] + [S \leftarrow I]) - 2\rho[I-I \leftarrow T] - 2g[I-I], \quad (8)$$
$$\frac{d}{dt}[S-T] = -\tau[T-S \leftarrow I] + \rho[S-I \leftarrow T]$$
$$\qquad + g[S-I] + g_T([T-T] - [S-T]),$$
$$\frac{d}{dt}[I-T] = \rho([I-I \leftarrow T] - [T-I \leftarrow T] - [I \leftarrow T])$$
$$\qquad + g([I-I] - [I-T]) - g_T[I-T],$$
$$\frac{d}{dt}[T-T] = 2\rho([T-I \leftarrow T] + [I \leftarrow T])$$
$$\qquad + g[I-T] - 2g_T[T-T].$$

We will also consider, for comparison, these equations closed using (4), and mean-field models.

Simulation

A complementary approach to pairwise models comes from individual-based, stochastic simulation where an explicit network is generated and dynamical processes are simulated using Monte Carlo methods.

In order to provide a good comparison between pairwise models and simulation, we generate explicit networks that are designed to introduce structure to the population along the lines that we have been considering, by introducing finite neighbourhood size and clustering, without introducing other significant structural features. This enables us to test results derived using pairwise equations against stochastic results. It also complements our general approach of looking at the implications of finite neighbourhood size and clustering *ceteris paribus*, as an aid to intuitive understanding of the impact of population structure on disease and intervention dynamics.

While other methods exist to generate networks with significant clustering coefficients, such as [20–22], and some special clustered networks have the attractive property of begin easily generated and analysed [23,24], we use simple rewiring methods that are easily described and whose implications for global network structure can be readily understood, but which limit us to a smaller region of network parameter space. Most importantly, we find that giant component sizes for networks generated using our methods typically exceed 99%.

Creation of a homogeneous random network. In order to create a homogeneous random network, we firstly generate a one-dimensional ring with k-th nearest neighbour links. Since we consider networks where n is even, we set $k=n/2$, and then make five cycles through every node i, and for each of that node's links $[i-j]$, swap with a random link $[k-l]$ as below,

$$
\begin{array}{ccccccc}
j & & l & & j & - - - & l \\
| & & | & \rightarrow & & & \\
i & & k & & i & - - - & k
\end{array}
\qquad (9)
$$

This generates a homogeneous random network free from dynamically relevant biases.

Increasing clustering. In order to increase the clustering coefficient for a network, whilst keeping degree distribution constant, we use a new rewiring method that we call the 'big V'. This means making the following network re-wiring for a 'V' of nodes $A-a-O-b-B$ as below,

$$
\begin{array}{c}
A \qquad\qquad\qquad B \\
\diagdown \qquad\qquad \diagup \\
a \qquad\qquad b \qquad \rightarrow \\
\diagdown \qquad \diagup \\
O
\end{array}
$$

$$
\begin{array}{c}
A \; - \; - \; - \; - \; - \; - \; - \; B \\
\\
a \; - \; - \; - \; b \qquad , \\
\diagdown \qquad \diagup \\
O
\end{array}
\qquad (10)
$$

provided the rewiring does not reduce the clustering coefficient overall. Clearly, such a rewiring does not modify the link distribution, but does increase the clustering coefficient. Empirically, we find that at low neighbourhood sizes, this method generates networks with clustering parameters up to $\phi=0.3$ before running out of possible rewirings. Whether alternative methods could yield larger values of ϕ without either a significant reduction in the giant component size or other dynamically relevant biases remains an interesting question, however the levels of clustering given by this rewiring are sufficient to demonstrate the qualitative epidemiological effects in which we are interested. Other recent work making use of this rewiring includes [25,26].

Stochastic dynamics. We simulate *SIS* dynamics with tracing on a network using a standard continuous-time algorithm [27]. The implementation of such algorithms, and the differences between them and discrete-time equivalents, in the context of epidemic models is discussed in [28, Chapter 6]. Since the two contact processes involved (infection and tracing) both involve the explicit network, our model is essentially individual based.

Parameterisation

For our baseline network parameters, we set $n=4$ to determine the effects of finite neighbourhood size and clustering. We also take the network size to be $N=10^5$ to produce little variability due to stochastic effects after the initial stages of an epidemic. Our main aim is to measure the effects of clustering, ϕ, and this is varied between 0 and 0.5. The recovery rate, g, can be formally set to 1 through non-dimensionalisation, and we set $g_T=10^3$ to achieve separation of timescales. Our epidemiological motivation for this separation is the expected difference in the time from infection to detection and treatment, and the time taken to notify sexual partners. For emerging respiratory infections, such a separation of timescales would, of course, not exist.

The other dynamical rates, τ and ρ are fixed indirectly. For the tracing rate, ρ, we vary the proportion of contacts successfully traced, $\varepsilon=\dfrac{\rho}{\rho+g_T}$, between 0 and 1, which then determines ρ. For the infection transmission rate, τ, we need methods for fitting to a given endemic equilibrium, in both stochastic and ODE contexts.

Pairwise transmission fitting. In the case of fitting to an endemic state, we solve the algebraic equations generated by setting

$$
\frac{d}{dt}[A]=\frac{d}{dt}[A-B]=0 \; , \quad [I]=I^*N \; , \qquad (11)
$$

in equations (8), giving a transmission rate τ^* that yields the default endemic equilibrium, $I^*=0.5$.

Stochastic transmission fitting. For computational efficiency, we use the following method to find the transmission rate τ^* needed to sustain a given endemic prevalence I^* at constant treatment rate g:

1. Each individual is set as infectious with probability I^* (and conversely, the probability of being set susceptible is $1-I^*$).
2. A random $[S-I]$ pair is chosen, and the susceptible individual is infected.
3. A random infectious individual is placed into the susceptible class.
4. Steps 2 and 3 are repeated until spatial structure is equilibrated, and then averages $\overline{[I]}$ and $\overline{[S-I]}$ of the number of infectious

individuals and susceptible-infectious pairs are taken over a further set of iterations of 2 and 3.

5. The transmission rate is then given by $\tau^* = g\dfrac{\overline{[I]}}{\overline{[S-I]}}$.

While this method is not simply described, it is accurate and, most importantly, computationally efficient.

Results

Dynamics in the absence of tracing

Figure 2 shows the comparison of stochastic simulation on networks of the type we have been considering with both mean-field *SIS*, standard pairwise, improved pairwise, and also the triplewise model of [29]. This demonstrates good agreement between simulation and network ODE models, but poor agreement with the mean-field model. The inclusion of the triplewise model shows that disagreements between pairwise models and simulation in the clustered network are largely due to higher order structure, however these effects are nowhere near as large as the differences between mean-field and pairwise models. Since triplewise models involve a massive increase in computational burden, we do not consider that in this case their use is justified.

The results of Figure 2 were obtained by fitting the improved pairwise model to a given endemic state, $I^* = 0.5$. The impact of this fitting on the transmission rate and number of $[I-I]$ pairs, while varying the clustering coefficient ϕ, is shown in Figure 2, panes C and D.

Impact of network structure on contact tracing

The need to incorporate network structure into models that involve contact tracing is shown by Figure 3. Panes A and B show predictions of prevalence over time for several models, which demonstrate that while both pairwise approaches are in good agreement with simulation, the failure of the mean-field model is dramatic—and similarly large failures can be observed in several other regions of parameter space.

For the case of a clustered network in Pane B, the agreement between pairwise models and simulation becomes slightly worse than for the unclustered network results of Pane A, with the improved pairwise model providing a closer fit. Most importantly, the improved pairwise model is in qualitative disagreement with simulation—while both mean-field and standard pairwise models predict a peak in infection before reaching the endemic state, which is not seen in either the improved pairwise model or simulation. We therefore use the results of Panes A and B to rule out the use of mean-field and standard pairwise models. This leaves the improved pairwise model, which we systematically compare to simulation in Panes C and D. Since both the improved pairwise model and simulation depend on underlying parameters in the same way, they form a complementary pair of approaches to the study of contact tracing in clustered populations. The only exception to this is the case of low prevalence of infection, where stochastic effects become important and the stochastic model predicts extinction at higher transmission rates than the pairwise model.
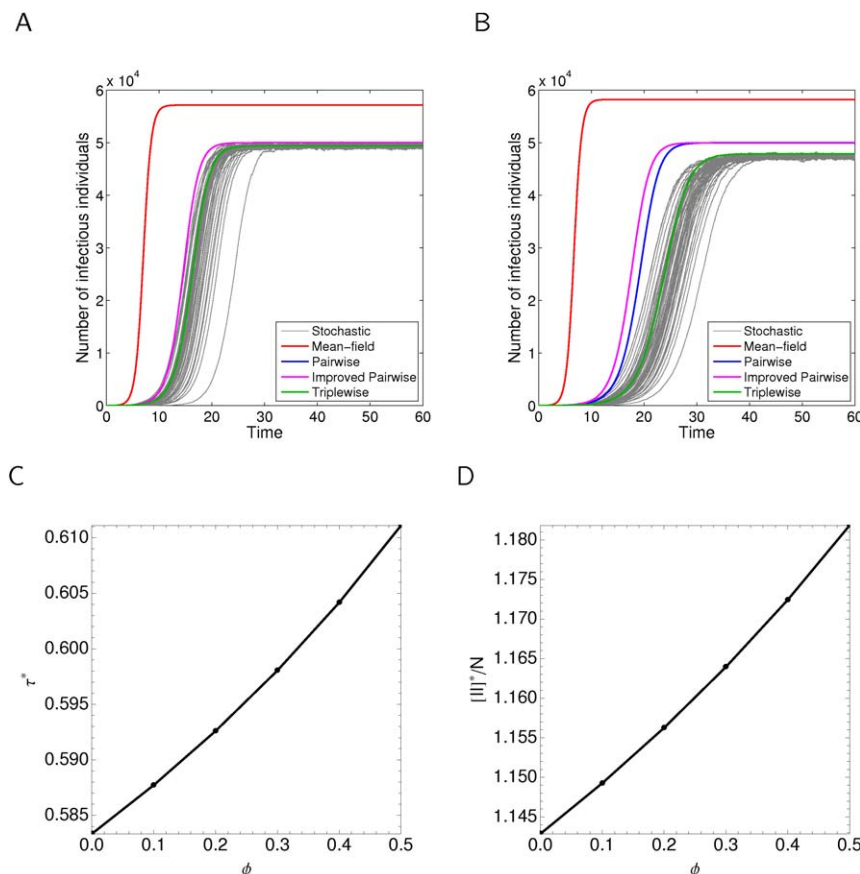

Figure 2. Features of *SIS* dynamics where the transmission rate τ is set in the improved pairwise model to give constant endemic state $I^* = 0.5$. Other parameters are set at their default values: $g = 1$, $n = 4$, $N = 10^5$. The clustering coefficient ϕ is set at either A: 0 or B: 0.3, and different ODE approaches are compared to simulation. Also shown are the values of C: τ^*, the transmission rate, and D: $[II]^*/N$, the relative weight of infectious-infectious pairs, at the endemic state as ϕ is varied while holding $I^* = 0.5$.
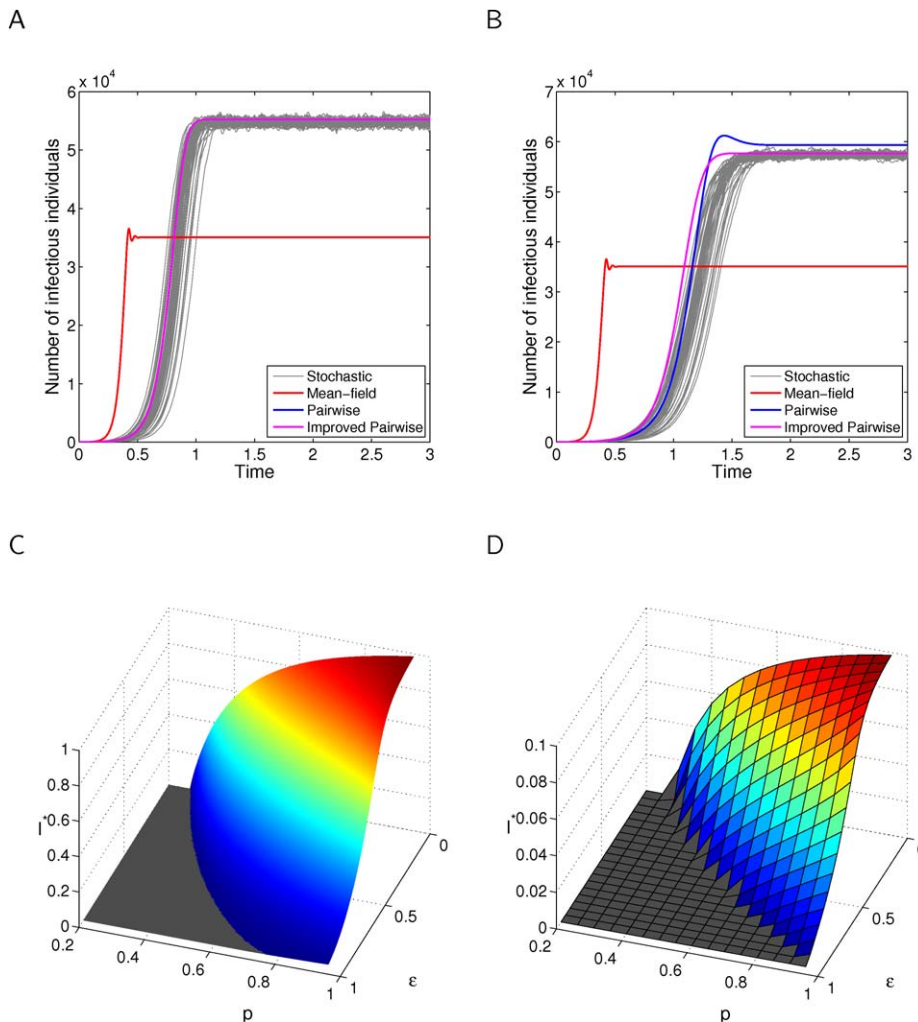doi:10.1371/journal.pcbi.1000721.g002

Figure 3. Comparison of infection curves for *SIS* dynamics with contact tracing. Clustering, ϕ is set to either A: 0 or B: 0.3, with contact tracing success, $\varepsilon = 0.4$ and infection rate, $\tau = 6$. Other parameters are set to their default values: $g = 1$, $g_T = 10^3$, $n = 4$, $N = 10^5$. The best agreement is between simulation and the improved pairwise model, with the mean-field approach qualitatively wrong. Sweeping over $p = \tau/(\tau + g)$ and ε for $\phi = 0.3$, the endemic states predicted by C: the improved pairwise model, and D: stochastic simulation on $N = 10^4$ nodes, are in good agreement except where prevalences are low.
doi:10.1371/journal.pcbi.1000721.g003

General results

We consider the effects of clustering on the efficacy of contact tracing using pairwise models by starting the system at the endemic state in the absence of any contact tracing. We then introduce tracing at a success probability ε, and allow the system to evolve away from the endemic state for 0.1 and one generations (time periods $1/(10g)$ and $1/g$ respectively, corresponding to policy evaluation after a number of months and a number of years for an endemic STI) and measure the numbers infectious. This gives the results in Figure 4, which show that clustering increases the efficacy of contact tracing at a given success rate at one infectious generation, but not at 0.1 generations, depending on the actual tracing success rate. Pane C shows this variety of responses, where clustering is more effective for large success rates at this small time—the very large rates require still smaller times to demonstrate this effect, since after 0.1 disease generations they have passed into the regime where clustering leads to less effective tracing.

The results shown in Panes C and D of Figure 2 provide a guide to intuition to explain these results. Clustering increases the

number of $[I - I]$ pairs present at a given endemic state, and contact tracing can be viewed as hyper-parasitism on the network of infected individuals. This means that clustering can be expected to enhance the efficacy of contact tracing by increasing the neighbourhood size of the infected sub-network. On the other hand, to explain a constant level of endemic infection as clustering is increased, a larger underlying rate of transmission must be present, which will undermine tracing as an individual left untouched by a wave of tracing will reinfect their immediate neighbourhood more quickly. Exactly which parameter choices allow either effect to dominate is not clear, except that lower levels of tracing success always cause clustering to increase the efficacy of tracing. Otherwise, it appears that the impact of clustering on contact tracing needs to be evaluated on a case-by-case basis.

Individual trajectories

To see the dynamics of infection that generate the results in Figure 4, we plot stochastic and improved pairwise temporal and parametric dynamics for two exemplar values of contact tracing
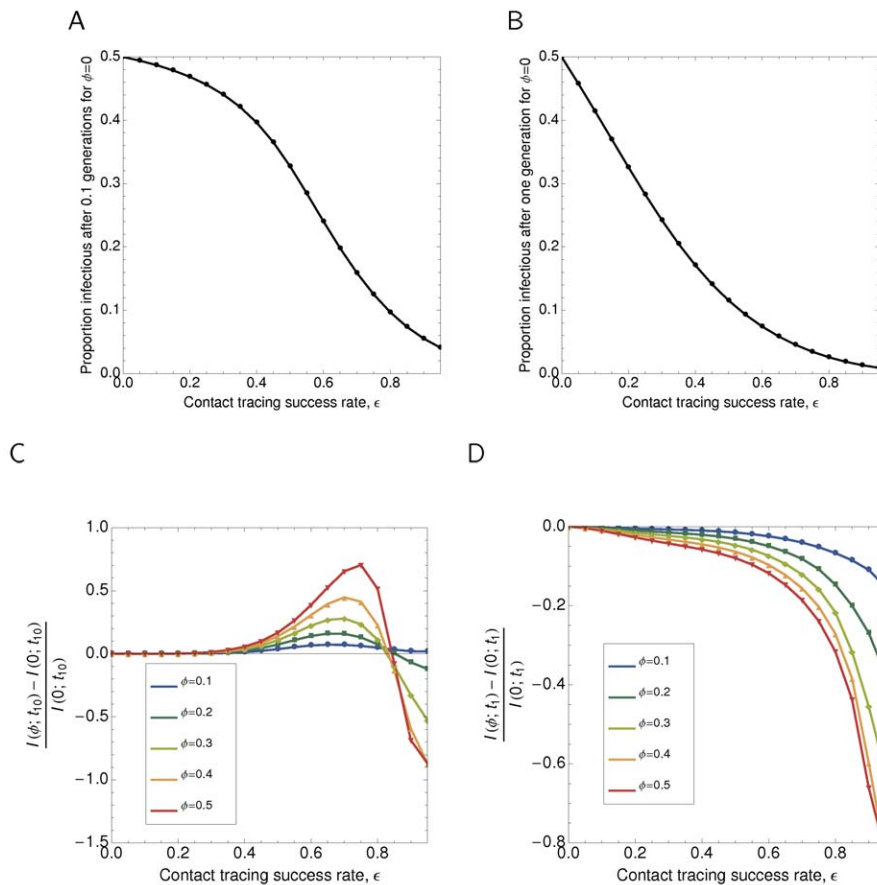
Figure 4. Impact of clustering on efficacy of contact tracing away from the equilibrium state with half of the population infectious.
The system is started in the endemic equilibrium, and then for A, B: $\phi = 0$ or C, D: $\phi > 0$, the level of infection is measured at times A, C: $t_{10} = 1/(10g)$ and B, D: $t_1 = 1/g$. Results are obtained in the improved pairwise model, with contact tracing success ε varied between 0 and 1. Other parameters set to defaults: $g = 1$, $g_T = 10^3$, $I^* = 0.5$.
doi:10.1371/journal.pcbi.1000721.g004

success, ε, in Figure 5. For $\varepsilon = 0.15$, we see in Pane A that the system settles over time to a different endemic state, and in Pane C that this involves consistently lower levels of infection in the clustered system than the unclustered system, meaning that clustering has enhanced the efficacy of contact tracing.

By contrast, for $\varepsilon = 0.65$, we see in Pane B that contact tracing drives infection to extinction, and from Pane D that this involves firstly higher levels of infection in the clustered system and then lower levels of infection for both the pairwise and stochastic models. We see a final reversal of the impact of clustering in Pane B, which is present in only the pairwise system: at longer times clustering again reduces the efficacy of contact tracing. At this point, stochastic variability in simulations has become highly significant and so we would not expect the two models to agree, since the pairwise equations should only hold in the limit where stochastic effects are negligible. To simulate in this regime would require extremely large population sizes, perhaps beyond what would ever be considered for realistic human scenarios.

Discussion

We have provided an intuitive and general framework in which to study the impact of network clustering on the efficacy of contact tracing in the control of infectious disease. This has produced three major results.

Firstly, the effects of contact tracing often cannot be accurately captured by mean-field models, necessitating a modelling approach that incorporates network structure.

Secondly, we have demonstrated that due to the increased number of infectious-infectious pairs seen in clustered networks at a given pathogen burden, contact tracing at a fixed, relatively low success rate will be more effective at larger levels of clustering than at the same fixed success rate without clustering.

Finally, we have demonstrated that this increased efficacy is not completely general, and is reversed for large tracing success rates at certain times. This demonstrates the need to be cautious in the consideration of the epidemiological effects of a network property as subtle as clustering—unfortunately it is not possible to obtain a general 'rule of thumb' for its impact.

Our approach has been to consider the impact of clustering on a network with fixed, finite neighbourhood size, in the absence of other known important dynamical effects such as risk structure and assortativity. The complexity of even our simplified problem provides justification for our approach, however it would be of significant interest to see how these quantities interact with each other. The full impact of higher order structure than triangles is also, as suggested by our stochastic results, likely to be important.

Another important difference may manifest itself if we were to consider a disease with long-lasting immunity, obeying *SIR* dynamics, or other compartment structure, including complex intervention strategies and comparable tracing and recovery timescales. Our
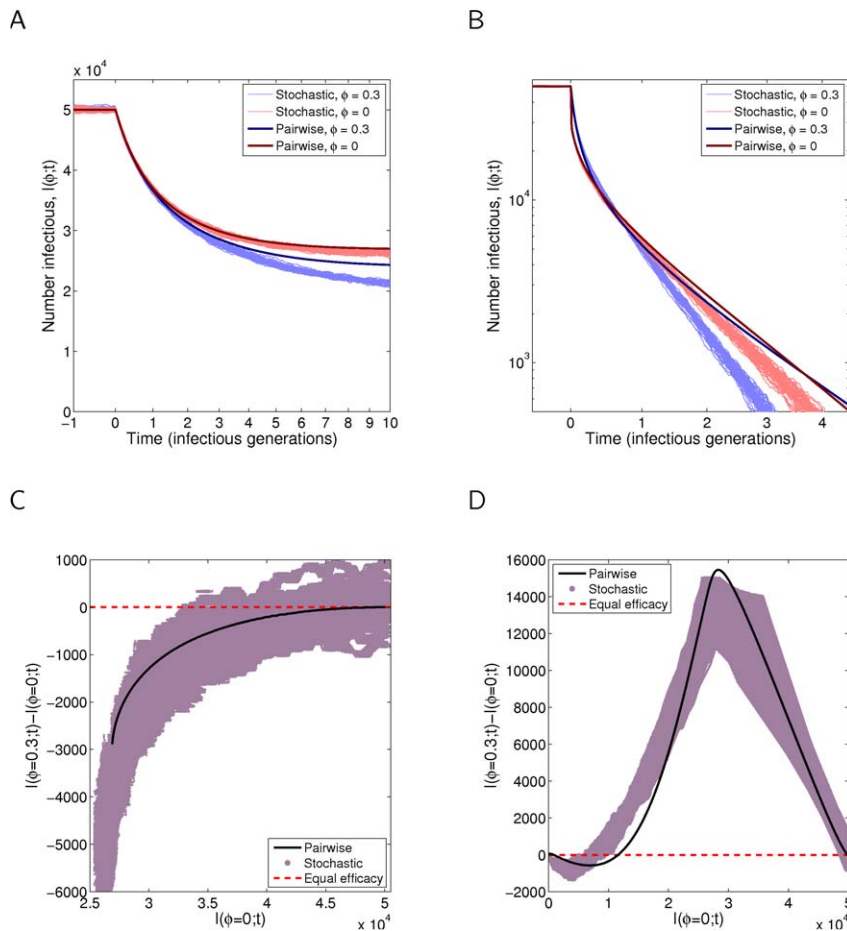
Figure 5. Temporal and parametric behaviour of contact tracing. The system is started in the endemic equilibrium, and then contact tracing is started with success rate ε equal to A, C: 0.15 or B, D: 0.65. Other parameters are at default values: $N = 10^5$, $g_T = 10^3$, $g = 1$, $I^* = 0.5$. The level of infection over time for different values of clustering ϕ are shown in Panes A and B, while in C and D the difference between clustered and unclustered is shown parametrically against the unclustered result—above the dotted line, clustering implies more efficacious tracing, and vice versa.
doi:10.1371/journal.pcbi.1000721.g005

preliminary work in this direction suggests that our novel result about clustering reducing contact tracing efficacy can be extremely significant in other contexts, however a full consideration of this would take us significantly beyond the aims of the present work.

Acknowledgments

Author Contributions

Conceived and designed the experiments: TH MJK. Performed the experiments: TH. Analyzed the data: TH. Contributed reagents/materials/analysis tools: TH. Wrote the paper: TH. Supervised the research: MJK.

References

1. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, et al. (2003) Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. Science 300: 1961–6.

2. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, et al. (2003) Transmission dynamics and control of severe acute respiratory syndrome. Science 300: 1966–70.

3. Ferguson NM, Donnelly CA, Anderson RM (2001) Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. Nature 413: 542–8.

4. Ferguson NM, Donnelly CA, Anderson RM (2001) The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. Science 292: 1155–60.

5. Keeling MJ, Woolhouse ME, Shaw DJ, Matthews L, Chase-Topping M, et al. (2001) Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science 294: 813–7.

6. Tildesley MJ, Savill NJ, Shaw DJ, Deardon R, Brooks SP, et al. (2006) Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. Nature 440: 83–6.

7. Riley S, Ferguson NM (2006) Smallpox transmission and control: Spatial dynamics in Great Britain. Proceedings of the National Academy of Sciences of the United States of America 103: 12637–12642.

8. Hall IM, Egan JR, Barrass I, Gani R, Leach S (2007) Comparison of smallpox outbreak control strategies using a spatial metapopulation model. Epidemiol Infect 135: 12.

9. Clarke J (1998) Contact tracing for chlamydia: data on effectiveness. International journal of STD & AIDS 9: 187–91.

10. FitzGerald MR, Thirlby D, Bedford CA (1998) The outcome of contact tracing for gonorrhoea in the United Kingdom. International journal of STD & AIDS 9: 657–60.

11. Golden MR, Hogben M, Handsfield HH, Lawrence JSS, Potterat JJ, et al. (2003) Partner notification for HIV and STD in the United States: low coverage for gonorrhea, chlamydial infection, and HIV. Sexually Transmitted Diseases 30: 490–6.

12. Müller J, Kretzschmar M, Dietz K (2000) Contact tracing in stochastic and deterministic epidemic models. Mathematical Biosciences 164: 39–64.

13. Huerta R, Tsimring LS (2002) Contact tracing and epidemics control in social networks. Physical Review E 66: 1–4.

14. Eames KTD, Keeling MJ (2003) Contact tracing and disease control. Proc Biol Sci 270: 2565–71.

15. Kiss IZ, Green DM, Kao RR (2007) The effect of network mixing patterns on epidemic dynamics and the efficacy of disease contact tracing. Journal of The Royal Society Interface 5: 791–799.

16. Ball F, O'Neill PD, Pike J (2007) Stochastic epidemic models in structured populations featuring dynamic vaccination and isolation. Journal of Applied Probability 44: 571–585.

17. Kiss IZ, Green DM, Kao RR (2005) Disease contact tracing in random and clustered networks. Proc Biol Sci 272: 1407–14.

18. Eames K, Keeling M (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. Proceedings of the National Academy of Sciences of the United States of America 99: 13330–13335.

19. Keeling MJ (1999) The effects of local spatial structure on epidemiological invasions. Proc Biol Sci 266: 859–67.

20. Volz E (2004) Random networks with tunable degree distribution and clustering. Physical Review E 70: 056115.

21. Serrano MA, Boguna M (2006) Percolation and epidemic thresholds in clustered networks. Physical Review Letters 97: 088701.

22. Eames KTD (2008) Modelling disease spread through random and regular contacts in clustered populations. Theoretical Population Biology 73: 104–11.

23. Newman MEJ (2009) Random graphs with clustering. Phys Rev Lett 103: 058701.

24. Miller JC (2009) Percolation and epidemics in random clustered networks. Phys Rev E 8: 020901(R).

25. Bansal S, Khandelwal S, Meyers LA (2009) Exploring biological network structure with clustered random networks. BMC Bioinformatics 10: 405.

26. Green DM, Kiss IZ (2010) Large-scale properties of clustered networks: Implications for disease dynamics. Journal of Biological Dynamics : In Press.

27. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem-US: 22.

28. Keeling MJ, Rohani P (2007) Modeling Infectious Diseases in Humans and Animals Princeton University Press.

29. House T, Davies G, Danon L, Keeling M (2009) A motif-based approach to network epidemics. Bulletin of Mathematical Biology.