AMSI
AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

National Collaboration in the Mathematical Sciences

Vaccination Methods for an Epidemic on a Multitype Random Population Network with Household Structure.

Ryan Heneghan

Introduction

Stochastic epidemic modelling has seen increasing attention from mathematicians in recent years, particularly in developing models that depart from the traditional assumption of a homogeneously mixed population of a single type of individual. This is in recognition of the fact that the homogenous model is unrealistic for all but the smallest of populations and that in almost all cases, individuals will respond to and spread an infection differently depending on their defining features (i.e., age, race, sociability et cetera). The reality of the complexity and size of modern population networks poses a unique challenge to public health professionals looking to find methods of distributing an often limited vaccine resource in a way that will minimise the final effects of an epidemic.

In this paper, I will use a population model developed by Ball & Sirl (2012) to attempt to find an optimal method of vaccination that reduces the final mortality of an epidemic. The model incorporates the heterogenous nature of a real world population network in three ways; the model incorporates several types of individuals on a random graph with household structure. A random graph structure ensures that each individual is only in contact with several others in the larger population, departing from the homogenous model which would be represented using a complete graph. Family structure in the graph is defined as the presence of complete subgraphs of a few members from the overall population, to model contact between individuals who have frequent contact with one another.

Individuals will be delineated into two classes based on their infectivity and the severity of the response to a possible infection. Class 2 will possess a comparatively low rate of infectiousness and a low rate of mortality, while Class 1 will have a high rate of mortality and a high rate of infectiousness. I will evaluate and compare three vaccination strategies:

1. Uniform random vaccine allocation across the entire population network, irrespective of class.

2. Vaccinate Class 1 uniformly at random and then, if there is sufficient vaccine to vaccinate all of class 1, also vaccinate Class 2 uniformly at random until vaccine supply is exhausted.

3. The opposite of 2.

Before evaluating these vaccine strategies a method of calculating the expected final size of the epidemic must be found. Subsequently, the rest of the paper will be organised as follows. In Section 1, I will describe the model and the conditions under which it will be analysed. In Section 2, I will evaluate the multitype analogue of the branching process approximation developed by Ball et. al (2009) that was used by Ball & Sirl (2012) to find the expected final size of a major outbreak. Finally in section 3, the results of a simulation of the vaccine methods will be evaluated and compared.

1 The Model.

There are two types of individuals in this population and m_η households of category η, where η denotes a houeshold with n_j individuals of type-j (for all $j \in \{1,2\}$). Let α denote the set of all household categories, so

$$\alpha = \{\eta = (n_1, n_2) \in \mathbb{Z}_+^2 : |\eta| \geq 1\},$$

where $|\eta| = n_1 + n_2$. Denote by $N_j = \sum_{\eta \in \alpha} n_j m_\eta$ the total number of type-j individuals and assume $N_j < \infty$ for all j, so it follows that $m_\eta > 0$ for only finitely many η. Let ρ_η denote the asymptotic proportion of households of type-η, given by the distribution $\rho = (\rho_\eta, \eta \in \alpha)$. Using ρ, we find that v_i, the proportion of individuals of type-i in the population, is given by $\sum_{\eta \in \alpha} n_i \rho_\eta / \sum_{\eta' \in \alpha} |\eta'| \rho_{\eta'}$ We consider each family as a connected subgraph with $|\eta|$ members and each individual in the population is a member of one family.

The global structure of the population model is given by a random graph where each individual has a degree distribution dependent on what class they are from. For each individual, we need to specify not just the total number of neighbours, but the total number of neighbours of each class. Let $\delta = (d_1, d_2)$ denote an individuals degree distribution, where d_i denotes the total number of neighbours from category-i. It's important to note that an individual's degree distribution gives the total number of adjacent individuals in the global population, not including those from its own household. We can consider each δ as the independent realisation of the random variable $\Delta_j = (D_j^1, D_j^2), j \in \{1,2\}$ with $\mathbb{P}(\Delta^{(j)} = \delta) = p_\delta^j$ for all $\delta \in \mathbb{Z}_+^2$. We assume that the mean number of neighbours of category-j for a type-i individual, $\mu_i^{(j)}$,is finite, i.e $\mu_j^{(i)} = \mathbb{E}[D_j^i] < \infty$ for all $i, j \in \{1,2\}$.

We can construct a realisation of the network of global contacts using a multitype version of the 'configuration model' (described in Durrett, (2006, Chapter 3)) as follows. Firstly, assign to each type-i individual a number of 'type $i \to j$ half-edges' according to independent realisations of the random variable $\Delta^i, i \in \{1,2\}$. Firstly, we pair the 'type $i \to i$ half-edges' together uniformly at random, with each pair of 'type $i \to i$ half-edges' forming an edge in the random graph describing the possible global contacts of type-i for a type-i individual. Then, for each pair i, j with $(i \neq j)$ and $i < j$, we take an $i \to j$ half-edge and pair it with a $j \to i$ half-edge chosen uniformly at random and repeat this process until one type of half edge is exhausted. To ensure that there are similiar numbers of type $i \to j$ and $j \to i$ half-edges, we enforce

the relation $v_1\mu_2^{(1)} = v_2\mu_1^{(2)}$, where v_i and $\mu_i^{(j)}$ denote the proportion of type-i individuals and the mean number of contacts of category-j for a type-i individual respectively.

There are a few imperfections in the random graph of global contacts that could arise using this construction. Namely, the issue of self-loops and parallel edges needs to be addressed. A self-loop would arise whenever two half-edges from the same individual connected, similiarly a parallel edge would occur when q half-edges from one individual connected to q half-edges of one other neighbour (where $q > 1$). If not properly controlled, self-loops and parallel edges may cause serious problems with the final analysis so it's important for the model to satisfy the conditions under which the probability of any randomly selected half-edge being involved in an imperfection converges in probability to 0 as $m \to \infty$. For this to occur the number of self loops and parallel edges would be required to not depend in any way on m as $m \to \infty$. Ball & Sirl (2012) show that, provided the degree distributions for each category of individual have finite second moments, the probability of a half-edge being involved in a self-loop or parallel edge converges to zero as $m \to \infty$.

Initially, all individuals would be susceptible except one infective. When an infective individual of type-i makes infectious contact with a susceptible individual that person will become infected at a rate dependent on both their category and whether they are within the household of the infected individual (a local contact) or from the greater global community (a global contact). A type-i infective makes infectious contact with any local type-j individual at the points of a Poisson process with rate $\lambda_{ij}^{(L)}$, similiarly with a type-j global contact this type-i infectious individual makes infectious contact at the points of a Poisson process with rate $\lambda_{ij}^{(G)}$. For convenience, we define the contact rate matrices as $\Lambda^{(L)} = (\lambda_{ij}^{(L)}, i, j \in \{1, 2\})$ and $\Lambda^{(G)} = (\lambda_{ij}^{(G)}, i, j \in \{1, 2\})$. Both $\lambda_{ij}^{(L)}$ and $\lambda_{ij}^{(G)}$ are per-pair infectious contact rates, so an infectious type-i individual from a family of category η and global degree distribution δ would make infectious contacts at the overall rate $\sum_{n_j \in \eta} n_j \lambda_{ij}^{(L)} + \sum_{d_j \in \delta} d_j \lambda_{ij}^{(G)}$. When an infective individual makes infectious contact with a susceptible individual of type-j, this newly infected individual will remain infectious for a period of time I determined by an exponential distribution with parameter γ, which is the same for all individuals. Upon leaving the infectious stage an infected type-i individual will either die with probability p_i or recover and become immune with probability q_i ($q_i = 1 - p_i$). So the total proportion of the population that dies is given by $z = p_1 z^{(1)} v_1 + p_2 z^{(2)} v_2$, where $z^{(i)}$ is the total proportion of type-i individuals who ultimately are infected. It's easy to see that the best vaccine method will be the one that first reduces the element of z which is initially

the largest. All Poisson processes and infectious periods are assumed to be mutually independent of each other and the population structure. The fact that the infectious periods and rates of infection follow mutually independent exponential and Poisson processes means that the epidemic could be represented as a continuous-time Markov chain.

1.1 An Example.

Ball & Sirl (2012) use parameters similiar to the following example for their simulation. We denote class 1 as children and class 2 as adults with the household category distribution (ρ_η) given in the table below (Table 1).

#adults	#children			
	0	1	2	3
0	-	0.00	0.00	0.00
1	0.205	0.04	0.04	0.02
2	0.195	0.15	0.25	0.10

This family distribution results in the proportions of individuals who are children (v_1) being $\frac{3}{5}$ and adults (v_2) being $\frac{2}{5}$. We take Δ^1 =(15,4) and Δ^2 =(6,10), so as an example this would mean that the average child would be in global contact with 15 other children and 4 adults. The infectious periods for each individual are independent, exponentially distributed random variables with mean parameter γ =10. We define the infectious contact rate matrices as

$$\Lambda^L = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{15} \end{pmatrix} \text{ and } \Lambda^G = \begin{pmatrix} \frac{2}{15} & \frac{1}{15} \\ \frac{1}{15} & \frac{1}{30} \end{pmatrix}.$$

So for example, the rate of infectious contact of a type-1 individual with a type-2 indivudal is given by an independent Poisson distribution with parameters $\frac{2}{3}$ if they are local contacts and $\frac{1}{15}$ if it is a type-2 global neighbour. We let the probabilities of dying if infected be $p_1 = 0.50$ for children and $p_2 = 0.20$ for adults. Figure 1 overleaf is a plot of the progression of an epidemic over a population of 500 families with the parameters just described.

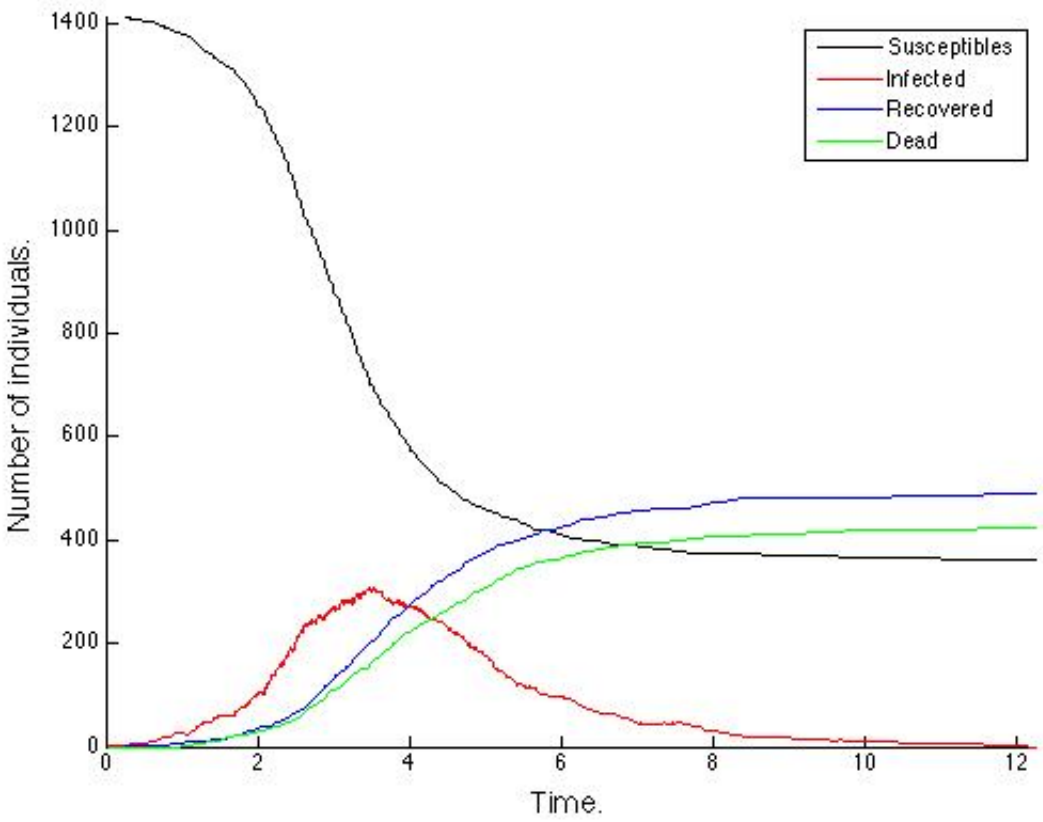


Figure 1: Progression of the Example Epidemic.

The peak of the epidemic occurs around 3.5 time intervals from the start of the epidemic, with just over 20% of the population is infected. After this, the number of new infections tapers off, before finally dying out around 12 time intervals from the start of the epidemic, with approximately 30% of the population dead and about 35% of the total population having become infected but surviving their infection and recovering.

2 Expected Relative Final Size of a Major Outbreak.

We now turn to calculating the expected proportion of individuals of a certain type that are ultimately infected by the epidemic ($z^{(i)}$). $z^{(i)}$ is equivalent to the probability that a randomly chosen type-i individual that is initially in an uninfected household is ultimately infected by a major outbreak. Ball & Sirl (2012) determine this probability by considering the size of the susceptibility set of a randomly chosen type-i individual.

2.1 Susceptibility Sets.

An individual's susceptibility set is a random set of individuals, dependent on the individual's infectious period and relevant independent Poisson processes associated with their infectivity, who would ultimately become infected were this individual of interest to become infected during an outbreak. More concretely, we construct a directed graph based on this set, in which a directed arc from i to j would be present when, if i were to become infected, it would make infectious contact with j. We can approximate the size of the susceptibility set of an individual by considering the total progeny of an appropriate approximating multitype branching process, denoted X. We can construct this branching process in terms of the households to which the individuals who comprise a random type-i individual's susceptibility set belong. The zeroth generation of X would comprise the household to which our individual of interest belongs and the first generation would be the households that contain an individual who makes global infectious contact with the individual's local susceptibility set (i.e, its household). Generalising, the $(n+1)^{th}$ generation would be the infected households that contain an individual who makes global infectious contact with a primary individual in a household from the n^{th} generation. Individuals who join the susceptibility set by making global infectious contact with a member of the existing susceptibility set are known as 'primary' contacts and those who join the set through within-household infectious contact are denoted as 'secondary' contacts. X is a 'backward process', as the infection spreads 'down' the generations (i.e, from n to $n-1$) before reaching the household of the individual of interest in the zeroth generation.

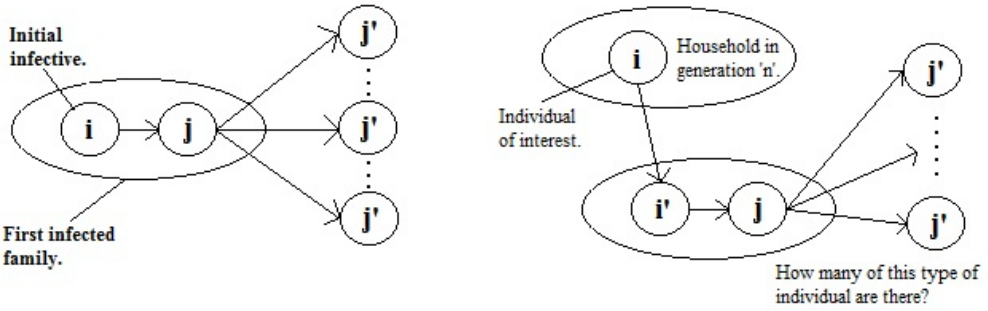
The full details of how the size of an individual's susceptibility set and the probability of its infection are very involved and what follows is a comparatively brief outline of the complete details. For the interested reader, the full argument in the case where all the individuals are from the same class can be found in section 6 of Ball *et al.*

(2009). Let $E^{(m_\eta)}$ denote the approximating branching process of an epidemic on a population network with m_η type-η households (for all $\eta \in \alpha$) and let $Z_t^{(m_\eta)}$ represent the number of infectious type-η households in generation t of $E^{(m_\eta)}$. Let $\Re_{i^*}^{(m_\eta)}$ denote the susceptibility set of a randomly selected type-i individual, denoted i^*, as described above. Now, we can consider the probability of i^* ultimately becoming infected when it is susceptible at generation t by stopping the construction of $E_{(m_\eta)}$ at generation t. This would leave $Z_t^m = \sum_{\eta\in\alpha} Z_t^{(m_\eta)}$ infectious half-edges unconnected (in this case, a complete edge in the construction of $E^{(m_\eta)}$ would consist of an infectious half edge from a household of generation n to a connecting to the half-edge of an individual in another uninfected household (generation $n+1$)). Returning to the construction of $\Re_{i^*}^{(m_\eta)}$, if at any point a half-edge from the growing susceptibility set of i^* is paired up with one of the Z_t^m infectious half-edges then i^* is ultimately infected.

Let $R^{(m_\eta)} = ((R_k^{(m_\eta)}), k = 0, 1, ...)$ describe the number of households in each generation of $\Re_{i^*}^{(m_\eta)}$ and let $X_{i^*}^{(m_\eta)}$ be a branching process associated with this susceptibility set that shares the same properties as X. Let $X^{(m_\eta)} = ((X_k^{(m_\eta)}), k = 0, 1, ...)$ denote the number of offspring in each generation of $X_{i^*}^{(m_\eta)}$. Let $\hat{R}^{(m_\eta)}$ denote the total number of households in $R^{(m_\eta)}$, further let $\hat{X}^{(m_\eta)}$ denote the total offspring of $X_{i^*}^{(m_\eta)}$. By definition, $\hat{R}^{(m_\eta)}$ and $\hat{X}^{(m_\eta)}$ have the same limiting distribution as $m \to \infty$. Now, if for any $k \in \mathbb{N}$, $\hat{R}^{(m_\eta)} < k$ then the probability that $\Re_{i^*}^{(m_\eta)}$ intersects with Z_t^m in a major outbreak goes to 0 as $m \to \infty$ for all t (since a major outbreak has at least $\log(m)$ infectious households (see Corollary 6.1 of Ball *et al.* (2009)). So as $m \to \infty$, the limiting probability that i^* is ultimately infected by a major outbreak is at most $\mathbb{P}(\hat{X}^{(m_\eta)} = \infty)$, the probability that the approximating branching process avoids extinction. In other words, since by definition $Z_t^{(m)}$ increases without bound as $m \to \infty$, $\hat{R}^{m_\eta}$ must not be bounded above (i.e, $\lim_{m\to\infty} \mathbb{P}(\hat{X}^{(m_\eta)} < k; k \in \mathbb{N}) = 0$).

2.2 The Branching Process.

To determine whether in the long term the branching process X goes extinct, we need to evaluate its offspring distribution. The offspring distribution needs to take into account the complexity of the population - both the family structure and the fact that there are multiple types of individuals. An individual who is part of a family in generation n of the branching process would have an offspring distribution that depends on its own category, the category of the individual who it makes global infectious contact with, the composition of that individuals local susceptibility set and the number of each type of individual that each member of that local susceptibility set is globally in infectious contact with. Ball & Sirl (2012) show that the random variable that describes this complexity can be denoted ${}_{ii'}\tilde{B} = ({}_{ii'}B_{jj'}, j, j \in \{1, 2\})$, where i is the category of the individual of interest in a household of generation n. ${}_{ii'}B_{jj'}$ is the number of type-j' individuals that make global infectious contact with an individual of type-j who is a member of the household of the type-i' individual who themselves have made global infectious contact with the individual of interest who is of type-i. Figure 2 overleaf gives a diagram of ${}_{ii'}\tilde{B}_{jj'}$. For the first generation, the random variable describing the offspring distribution is denoted ${}_{i}\tilde{B} = ({}_{i}B_{jj'}, j, j' \in 1, 2)$, where ${}_{i}B_{jj'}$ is the number of type-j' individuals who make global contact with a type-j individual who is in the local susceptibility set of the initial type-i infective. It is assumed that, as $m \to \infty$, during the early growth of the susceptibility set, the type-j' individuals who are joining the set are in households that are previously unassociated with the susceptibility set almost surely. This ensures the branching property of the susceptibility set, which is essential if we are to approximate its growth using the backward process described in section 2.1.



(a) Initial Generation (${}_{i}B_{jj'}$). (b) Subsequent Generations (${}_{ii'}B_{jj'}$).

Figure 2: Diagrams of ${}_{i}B_{jj'}$ and ${}_{ii'}Bjj'$.

Now that the random variables that describe the offspring distributions for the branching process have been detailled, we need to determine the probability that these variables equal 0 as $m \to \infty$ (i.e, the probability of the branching process ultimately going extinct). For any single-type, homogenous population where Z_n is the size of generation n, the probability of the ultimate extinction of the branching process as $n \to \infty$ can be found by finding ϵ, the smallest non-negative root of the processes probability generating function (PGF) $G(s)$, (i.e, ϵ is the smallest $s \in [0,1]$ for which $G(s) = s$). For full details and a proof of this theorem, see Theorem 5.4.5 of Grimmett & Stirzaker (2001). Ball & Sirl (2012) use a multitype analogue of this theorem on the PGFs of ${}_{ii'}\tilde{B}$ and ${}_{i}B$, which are denoted

$$\tilde{H}_{ii'}(\mathbf{s}) = f_{{}_{ii'}\tilde{B}}(\mathbf{s}) = \sum_{\mathbf{x}\in R} \mathbb{P}({}_{ii'}\tilde{B} = \mathbf{x})\mathbf{s}^{\mathbf{x}} = \mathbb{E}[\mathbf{s}^{{}_{ii'}\tilde{B}}],$$

and

$$H_i(\mathbf{s}) = f_{{}_{i}\tilde{B}}(\mathbf{s}) = \sum_{\mathbf{x}\in R} \mathbb{P}({}_{i}\tilde{B} = \mathbf{x})\mathbf{s}^{\mathbf{x}} = \mathbb{E}[\mathbf{s}^{{}_{i}\tilde{B}}],$$

where $R \in \mathbb{Z}_+^4$, $s \in [0,1]^4$, $(i,i') \in \{1,2\}^2$ and $i \in \{1,2\}$. The full derivation of these PGFs can be found in section 3.2 of Ball & Sirl (2012). If we write $\tilde{\mathbf{H}}(\mathbf{s}) = (\tilde{H}_{ii'}(\mathbf{s}), i, i' = 1,2)$, the asymptotic probability that a type-i individual not in the initially infected household is ultimately infected in the epidemic given by $z^{(i)} = 1 - H_i(\mathbf{h})$, where $\mathbf{h}$ is the smallest solution of $\tilde{\mathbf{H}}(\mathbf{s}) = \mathbf{s}$ ($\mathbf{s} \in [0,1]^4$). So, the total probability of any randomly chosen individual not in the initial infected household becoming ultimately infected is given by $z = \sum_{j\in\{1,2\}} v_j z^{(j)}$. As was mentioned at the beginning of section 2, $z^{(j)}$ is also the total proportion of type-j individuals ultimately infected in the epidemic, so z is the total proportion of the population infected.

3 Testing the Vaccination Strategies.

We now turn to comparing the effects of the vaccine strategies detailed in the introduction on the overall mortality of an epidemic. Before doing so, it's essential that the effect of the vaccine be explained. In this paper, we are examing only a nonrandom vaccine, meaning vaccinated individuals of the same type all have the same response to the vaccine. We say that for each $i \in \{1, 2\}$, a vaccinated type-i individual has relative susceptibility $a_i \in [0, 1]$ (compared to a non-vaccinated type-i individual) and, if it were to become infected, a relative infectivity $b_i \in [0, 1]$. So, the rate of each Poisson process associated with a particular type-i individual becoming infected is multplied by a_i, which can be interpreted as the vaccinee 'repelling' each incoming infectious contact with probability $1-a_i$. Similiarly, if the vaccinated type-i individual were to become infected, the rates at which it makes infectious contact would be all multuplied by b_i. For our purposes in this paper, we will take $a_i = 0$ for all $i \in \{1, 2\}$. So the vaccine immunises individuals perfectly across all categories. We can set b_i to any value, as a vaccinated individual can never become infected and so cannot be part of the spread of the epidemic. For the sake of continuity, the parameters for the example epidemic in section 1.1 will be used here.

The vaccination strategies are simulated as follows. We begin with a population of susceptibles with 0% vaccinated and simulate the epidemic, we then increase the number of vaccinees one percentage point at a time by removing them from the population of susceptibles before the epidemic begins in a fashion that depends on the vaccination strategy being simulated. When the epidemic has run its course, we find the total proportion of the population that has died for this run of the simulation. Figure 3 overleaf was obtained in this manner, by running the simulation over a population of 300 families (approximately 900 individuals) where each percentage point was simulated 50 times for each of the three vaccine strategies. The lines tracking the end result for each strategy are found by calculating the mean of the nonzero results over the 50 simulations for each percentage point. In total, this required 15,150 simulations, which took approximately 20 hours on a desktop computer.

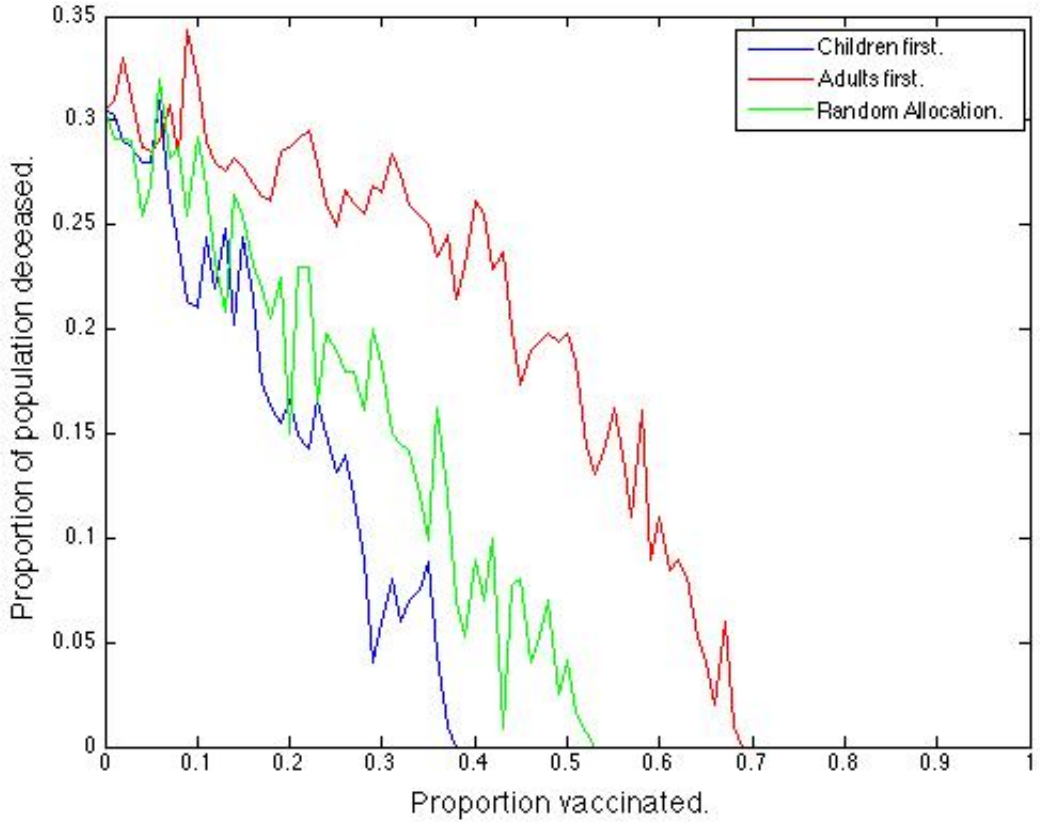


Figure 3: The Vaccination Strategies.

There is a fair amount of variability in the lines for each of the vaccination strategies. This could be remedied by running more simulations over a larger population; Ball & Sirl (2012) simulate the epidemic on a population of 500 families, with 1,000 simulations for each percentage point of each of the vaccination strategies and obtain a much smoother plot (see Figure 2, section 5 of Ball & Sirl (2012)). However Figure 3 does share a similiar shape to Ball & Sirl's result and the points at which the proportion of people vaccinated prevent the epidemic taking off for each vaccination strategy are almost the same.

Immediately it is apparent that vaccinating the children first is the better strategy in this situation. The 'adults first' strategy requires almost 80% more vaccine to achieve a similiar result, with 68% vaccination required, compared to the 'children first' strategy, with approximately 37% coverage required. It's worth remembering also that children make up 40% of the population, this indicates that children are the ones primarily responsible for the spread of the infection and that with the children almost entirely vaccinated the epidemic will not spread amongst a population almost totally made up of adults. However the 'adults first' strategy indicates that one needs to vaccinate all the adults first, and then about one fifth of the children before the epidemic cannot spread on the remaining susceptibles, which means that we cannot dismiss adults as not playing an important role in the spread of the epidemic.

4 Conclusion & Acknowledgements.

There is a multitude of possible extensions that could be applied to this model and the vaccination strategies. For example, a quite effective vaccination strategy involves identifying an infective and effectively 'quarantining' them by vaccinating their local and global contacts. There is also the difficulty of the theory of the population model and epidemic meeting real world observations - most of the time the parameters of an epidemic are unknown and must be estimated after the fact, for example the infectious period distribution of different types of individuals or the rates of the spread of the infection amongst different classes.

I'd like to thank Phil Pollett, my supervisor at UQ for being so generous with his time and knowledge and for proposing the idea of applying for the AMSI scholarship in the first place. I'd also like to thank AMSI and CSIRO for the scholarship and the Big Day In. The experience I gained through this project has been invaluable in shedding light on what I am capable of and possible directions for my honours year and career beyond.

References.

- BALL, F. & LYNE, O.D. (2001). 'Stochastic multitype SIR epidemics among a population partitioned into households.' *Adv. Appl. Prob.* vol. 33, pp. 99-123.
- BALL, F. & SIRL, D. (2012). 'An SIR epidemic model on a population with random network and household structure, and several types of individuals.' *Adv. Appl. Prob.* vol. 44, pp. 63-86.
- DURRET, R. (2007). *Random Graph Dynamics*, Cambridge University Press, Cambridge, England.
- GRIMMETT, G. & STIRZAKER, D. (2001). *Probability and Random Processes*, 3rd edn, Oxford University Press, Oxford, Engalnd.