Containing Pandemic Influenza at the Source:

Supporting Online Material

Ira M. Longini, Jr.[1], Azhar Nizam[1], Shufu Xu[1],

Kumnuan Ungchusak[2], Wanna Hanshaoworakul[2],

Derek A.T. Cummings[3], M. Elizabeth Halloran[1]

July 12, 2005

[1]Department of Biostatistics, The Rollins School of Public Health, Emory University, 1518 Clifton Road, N.E., Atlanta, GA, USA

[2]Ministry of Public Health, Nonthaburi, Thailand

[3]Department of International Health, The Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

1 The Population

The model population for a rural South East (SE) Asian population of 500,000 people is based on Thai census data (1), demographic information (2) and social network data (3). This model population is distributed across a 12.5 km $\times$ 12.5 km = 5,625 km^2, yielding a population density of ~89/km^2. This is the population density of rural Thailand (1). The whole population is partitioned into 36 localities each of size ~14,000 people in an area of 12.5 km $\times$ 12.5 km = ~156/km^2 (Fig. S1). Each locality consists of ~28 villages, each with ~138 households and ~500 people (Fig. S2). Villages are clustered and households are clustered within villages as described below. The household size and age distributions of the model population are given in tables S1 and S2, respectively, and match the distribution of rural Thailand (1). We populated the households with adults and children of various ages to match these distributions. With the exception of households, mixing groups are sized to reflect the potential number of people with whom there could be contact sufficient to transmit infection. For example, we assign working people to work groups of approximately 21 people. Although some actual work groups may be much larger than 21 people, we assume that workers do not make contact with every one in the larger work group but with a smaller group of people with whom they directly interact. People are assigned to workplaces according to a function of distance traveled from home to work that is described below. Small children, aged six months to four years mix in village level small and larger groups, with average sizes four and 22 children, respectively. Children 5-10 years of age were assigned to local elementary schools of average size 117 children in a mixing group, such that each school spanned two villages (3). Children 11-14 years of age were assigned to lower secondary schools of average size 95 based on the distance function. In the model population, 17% of children in the 11-14 year age group do not attend

school (1) and are assigned to the work force. Children 15-17 years of age were assigned to upper secondary schools of average size 69 based on the distance function. In the model population, 42% of children in the 15-17 year age group did not attend school (1) and are assigned to the work force.

People are assigned to daily work place mixing groups according to data on the distance that workers traveled to work in Thailand (Fig. S3) (2,4). These distance data indicate very localized movement of the population. For each locality, we randomly assign the workers in that locality to their work locality according to the distribution shown in Fig. S3. Fig. S4 gives an example of such an assignment for people located in the most southwesterly locality of the 500,000 person population. People assigned to the workplace mixing groups are 82% of the adults and the percentages of school children given above.

All people in the population travel to and mix per day in two social groups of average size 100. These groups represent the mixing groups where causal, untraceable contacts take place such as temples, markets, shops and other places where people mix casually. We assign people to these groups according to the distance distribution described above.

Although nosocomial spread has not been important in past influenza pandemics, we include one centrally located 40-bed community hospital that serves the 36 localities (5). People with influenza illness are hospitalized according the age-specific rates from Hong Kong physicians (Table S3) (6). We also created one influenza case holding center of size 100 in each of the 36 localities. If the hospital becomes full, an influenza case referred to hospitalization is assigned to the closest holding center that still has room. If the hospital and all 36 holding centers are full, then the case is returned to his or her home. There are 40 hospital staff. Each person in the population has a daily probability 10^{-3} of going to the hospital for a reason unrelated to influenza.

2 Contact process

For transmission of influenza to occur, people must make contact sufficient to transmit. We define the mixing group-specific probability of sufficient contact per day as c. Also, c can be a function of the age groups making contact. Given sufficient contact is made between a fully infectious and fully susceptible person (e.g., without vaccine or antivirals), transmission occurs with probability x, where we define x as the per contact transmission probability. Note that x does not vary by age or mixing group. Then, the probability that a fully infectious person infects a fully susceptible person per day in a mixing group is $p = cx$. We determine the values of c and x to give the baseline age specific attack rates given in table S4. In addition, the values p for the household mixing groups are set to give household secondary attack rates that are consistent with past influenza epidemics (7-10). Table S5 gives the values of c used for the baseline influenza epidemic. For example, the probability that two children make sufficient contact per day to transmit influenza in a household is 0.6, while that probability between a child and an adult is 0.3. If there are n people in a mixing group, then $c(n-1)$ is the average number of sufficient contacts that a person makes with others in the mixing group. This average number is age stratified in the households. Then $c(n-1)$ is the average daily contact degree for the number of contacts in a mixing group. Table S6 gives the average contact degree for the various mixing groups in the population. The highest degree is 7.2 among elementary school children. Thus, an elementary school child makes an average of 7.2 contacts sufficient to transmit per day with other elementary school children in that mixing group. We have made the contact structure as consistent as we can with that of rural Thailand (3).

To analyze the connectivity among the 500,000 people in our simulation population, we generated realizations of a person-to-person contact graph based on the sufficient contact probabilities

in table S5. A weighted random graph was constructed from the contact probabilities and the assignment of people to the different mixing groups based on the social network. We then generated realizations of the random graph by sampling from the weighted person-to-person graph (11). The probability distributions for the contact degree, clustering, and shortest path were generated from 40 realizations. Fig. S5 shows the degree distribution for the entire population. The mean sufficient contact degree is 4.6, however the distribution has a long right tail indicating that a few people make a large number of contacts. The average clustering coefficient for the contact network is 0.2 and the mean shortest path between randomly selected people is 10.6 links. This indicates that the contact network is quite clustered with relatively short links between these clusters that would characterize it as a small world network (11-12). Figs. S6- S8 give the contectivity patterns for three selected people in the contact network. Fig. S9 gives the average number of sufficient contacts made per day between an average person in each age group. These numbers are averaged across all the mixing groups with whom people in a particular age group mix. Again, we see that the largest number of daily sufficient contacts is made among elementary school children, where the average elementary school child makes an average of 7.5 contacts with other elementary school children.

3 Transmission process

The probability of infection for each susceptible individual each day is based on the transmission probabilities for each potentially infectious contact, $p = cx$. As an example, consider the simplest case that no one is vaccinated or using antiviral agents. An elementary school child is exposed to the number of child and adult infectives in his household, I_{hc} and I_{ha}, his neighborhood cluster, I_{nc}

and I_{na}, his school I_{es}, and the two social groups he or she mixes in, I_{s1} and I_{s2}, with corresponding transmission probabilities for each contact of p_{hcc} (child to child), p_{hac} (adult to child), p_{ncc}, p_{nac}, p_{es}, p_{s1} and p_{s2}, respectively. The probability P for that child to become infected on that day is

$$P = 1 - (1 - p_{hcc})^{I_{hc}}(1 - p_{hac})^{I_{ha}}(1 - p_{ncc})^{I_{nc}}(1 - p_{nac})^{I_{na}}(1 - p_{es})^{I_{es}}(1 - p_{s1})^{I_{s1}}(1 - p_{s2})^{I_{s2}}.$$

A Bernoulli trial is conducted by generating a uniform [0,1] random number. If the number is less than P, the child becomes infected and enters the latent phase of infection. For infected people, the source of infection is determined by conducting further Bernoulli trails for each mixing group.

If exposed people have been given antiviral agents, the transmission probabilities are multiplied by θ, the relative susceptibility, where protective efficacy $\text{AVE}_S = 1 - \theta$. If an infected person is using an antiviral, then the transmission probability from that infected person to a susceptible not using an antiviral is multiplied by ϕ, the relative infectiousness of infectives. The antiviral efficacy for infectiousness is $\text{AVE}_I = 1 - \phi$. If a person using an antiviral agent is infected, then the probability that he will become ill is multiplied by ψ, the relative probability of illness given infection. Thus, the antiviral efficacy for illness given infection is $\text{AVE}_D = 1 - \psi$. In addition, if a person on antiviral agents does become ill, then his duration of illness is one day less than if he had not taken an antiviral agent. If a person takes an antiviral agent after he is infected, then the AVE_I and AVE_D apply as above, from the time such use begins. We assume that $\text{AVE}_S = 0.30$, $\text{AVE}_I = 0.62$, and $\text{AVE}_D = 0.60$. For vaccination, we use arguments similar to those above to define vaccine efficacy for susceptibility, VE_S,and vaccine efficacy for infectiousness, VE_I. We assume that $\text{VE}_S = 0.30$, $\text{VE}_I = 0.50$.

4 Antiviral Efficacy

We use current estimates of antiviral efficacy (AVE) of oseltamivir (13-15). For prophylactic use, we assume that the antiviral efficacy for susceptibility to infection (AVE_S) is 0.30. For both therapeutic and prophylactic use, we assume that the efficacy for symptomatic disease given infection (AVE_D) is 0.60, and the antiviral efficacy for symptomatic disease (AVE_{SD}) is $\text{AVE}_{SD} = 1 - (1 - \text{AVE}_S)(1 - \text{AVE}_D) = 0.72$. We assume that the antiviral efficacy for infectiousness (AVE_I) is 0.62 (16). The effect of antiviral agent on people who take it while latent or incubating is the same as the therapeutic effect. A single course of oseltamivir consists of 10 tablets, enough for five days of treatment or 10 days of prophylaxis.

5 Basic reproductive number (R_0)

The basic reproductive number, R_0, is defined as the average number of secondary infections produced by a typical infected person in a fully susceptible population [17]. For a heterogeneous population, it is the average of all the secondary cases that this typical initial infective would infect over all the mixing groups that he or she mixes in. If these mixing groups do not overlap, *i.e.*, form a partition, and the model is deterministic, then R_0 is the dominant eigenvalue of the next generation matrix (17-18). In a stochastic model, the calculation is more involved and an approximation is analytically tractable only when the mixing groups form a partition (19). Our model has overlapping mixing groups and is stochastic, thus, the R_0 must be empirically calculated. We do this directly from the definition of R_0. We select a typical initial infective by randomly selecting a person from the population. To calculate R_0, we assumed a scenario in which one randomly chosen, unvaccinated infected person is seeded into a population where everyone else's ability to

transmit is 0. We then count the number of secondary infections. This is repeated 1000 times. In this way, we generate the whole distribution of secondary cases due to a randomly selected infected person. The random selection of the initial infective insures proper sampling from all possible initial infective types throughout the contact network. The mean of this distribution is R_0. We have found that this method generates $R_0$with the correct threshold condition (20).

6 Baseline epidemics

Based on the sufficient contact probabilities given in table S4 and a transmission probability of $x = 0.1,$ the baseline epidemic with age specific illness attack rates shown in Fig. 1.b. was generated with $R_0 = 1.4$ (see Fig. S10). From Fig. S10, we see that there is considerable variation in the number of secondary cases. Nearly 30% of the time, there are no secondary cases due to the single introduction. There is a small probability that a single introduction directly infected 13 other people. We produce the other levels of R_0 shown in Fig. 1.b, by varying the value of the transmission probability x. The baseline epidemic produces sources of infection shown in table S7. Thus, close contact groups account for 89% of the infections and casual contacts just 11%.

7 Sensitivity analyses

Fig.s S11-S15 show the sensitivity analysis of the effectiveness of the different interventions at different levels of implementation and delay. At $R_0 \leq 1.4$, both TAP or GTAP alone are effective only at the 70% level or higher. TAP is somewhat more effective than GTAP if $R_0 \leq 1.4$, but GTAP is more effective than TAP if $R_0 \geq 1.7$. In addition, TAP uses many fewer courses of oseltamivir

than GTAP if $R_0 \leq 1.4$, but GTAP uses somewhat fewer courses than TAP if $R_0 \geq 1.7$ (see Table 1). At higher values of R_0, household quarantine must reach the 70% level to be effective. From figure S14, 90% GTAP and 80% TAP become less effective when the intervention starts 28 days or more after the detection of the first symptomatic case, when there is an average of 85 cases already. Quarantine at the 70% level becomes less effective 42 days (average 313 cases) after detection of the first case, even with the addition of TAP. All other interventions in combination with pre-vaccination would be effective even 56 days (average 894 cases) after the detection of the first case.

In these simulations, we assumed that the initiation of TAP would begin one day after the first symptomatic illness, *i.e.*, the index case, in the close contact mixing groups. However, it may not be practical to get antiviral agents to exposed people so quickly. We carried out a sensitivity analysis for delays ranging from two - five days after detection of an index case, with 80% TAP. With a delay of up to 2 days, substantial reduction in the number of cases is still achieved, but with delays of 3-5 days, there is less benefit (Fig. S15).

Fig. S16 shows that the effectiveness of TAP and GTAP are moderately sensitive to variation in the AVE$_S$. The AVE$_S$ should be 0.5 or higher for the either TAP or GTAP to be effective. Fig. S17 shows that the effectiveness of TAP and GTAP are not very sensitive to variation in the AVE$_D$. Fig. S18 shows that the effectiveness of TAP and GTAP are quite sensitive to variation in the AVE$_I$. The AVE$_I$ should be 0.5 or higher for the either TAP or GTAP to be effective.

Fig. S19 shows that the effectiveness of 80% TAP with 70% pre-vaccination is sensitive to variation in the VE$_I$. However, even at level of VE$_I$ as low as 0.1, the epidemic is still well contained.

8 Epidemic movie descriptions (see the attached CD)

- **Movie 1**: Simulation of two epidemics of a newly emergent influenza strain on the geographic space of 500,000 people in rural SE Asia. The no intervention and interventions epidemics are single stochastic realizations when $R_0 = 1.4$. The intervention is 80% TAP initiated 14 days after the first detected case. Both the no intervention and intervention epidemics are started with the same initially infected person, and the starting random seed is the same for both stochastic realizations. Thus, the intervention shows what would happen under the same circumstances that produced a large epidemic with no intervention. The left two panels display the two simulated epidemic curves (i.e., number of new cases per day), and the right two panels display the corresponding epidemics on the maps. Dots representing people are drawn on the maps, with yellow dots for infected people and blue dots for recovered or dead people.

- **Movie 2**: The simulated epidemic curves in each locality from a single realization of an epidemic of a newly emergent influenza strain with no intervention when $R_0 = 1.4$. This is the same realization as given in the top panels of movie 1.

- **Movie 3**: The simulated epidemic curves in each locality from a single realization of an epidemic a newly emergent influenza strain with the intervention is 80% TAP initiated 14 days after the first detected case when $R_0 = 1.4$. This is the same realization as given in the bottom panels of the movie 1.

References

[1] National Stistical Office. Population and housing census 2000. http:/WWW.nso.go.th, accessed 19/11/05.

[2] P. Guest, A. Chamraathrithirong, K. Arhavanitkul, and N. Piriyathamwong. *Asian Pacific Migrat. J.*, 3:531, 1994.

[3] Faust K., Entwisle B., Rindfuss R.R., Walsh S.J., and Sawangdee Y. *Soc. Networks*, 28:311, 1999.

[4] A. Chamraathrithirong et al. *National Migration Survey of Thailand*. Institute for Population and Social Research, Mahidol Univeristy, Bangkok, 1995.

[5] Thailand health profile 1999-2000. pages 279–291. http:/WWW.moph.go.th/ops/thealth 44/ Retrieved 15/3/05.

[6] Fitzner K.A., McGhee S.M., Hedley S.M., and Shortridge K.F. *Hong Kong Med. J.*, 5:88, 1999.

[7] CL Addy, IM Longini, and MS Haber. *Biometrics*, 47:961, 1991.

[8] IM Longini and JS Koopman. *Biometrics*, 38:115, 1982.

[9] IM Longini, JS Koopman, M Haber, and GA Cotsonis. *Am J Epidemiol*, 128:845, 1988.

[10] S. Cauchemez, F. Carrat, C. Viboud, A.J. Valleron, and P.Y. Boëlle. *Stat. Med.*, 23:3469, 2004.

[11] M.E.J. Newman. *SIAM Review*, 45:167, 2003.

[12] S. Eubank et al. *Nature*, 429:180, 2004.

[13] FG Hayden and FY Aoki. In V Yu, T Meigan, and S Barriere, editors, *in Antimicrobial Therapy and Vaccines*, pages 1344–1365. Williams and Wilkins, Baltimore, MD, 1999.

[14] FG Hayden et al. *New England Journal of Medicine*, 343:1282, 2000.

[15] R Welliver et al. *JAMA*, 285:748, 2001.

[16] Y Yang and IM Longini. New methods for the estimation of influenza antiviral agent efficacy. In *in Joint Statistical Meetings*, San Francisco, August 3 - 7 2003.

[17] O Diekmann, JAP Heesterbeek, and JAJ Metz. *J Math Biol*, 28:365, 1990.

[18] AN Hill and IM Longini. *Math. Biosci.*, 181:85, 2003.

[19] T Britton. *J. Roy. Statist. Soc. B*, 63:705, 2001.

[20] IM Longini, ME Halloran, A Nizam, and Y Yang. *Am J Epidemiol*, 159:623, 2004.

FIGURE CAPTIONS

1. Fig. S1: Schematic of the model population for a rural South East (SE) Asian population of 500,000 people based on Thai statistics. The area covered is 75 km $\times$ 75 km = 5,625 km^2. Thus, the population density is ~89/km^2. Shown is the partition of the population into 36 localities each of size ~14,000 people in an area of 12.5 km $\times$ 12.5 km = ~156/km^2.

   Fig. S2: Schematic describing one of the localities consisting of ~28 villages, each of size ~138 households and ~500 people. The various contact mixing groups are clustered within the locality according to statistics on the social structure of rural Thai populations.

   Fig. S3: Histogram showing data on the percentage of people in rural Thailand that travel various distances to work (2), schools, and social groups. Travel is highly localized with 90% of the trips staying within 15 km. This roughly corresponds to the proportion of people who would not travel outside of their locality.

   Fig. S4: Schematic showing how the people in the most southwesterly locality are assigned to workplaces, schools and social groups in the model. People in other localities are assigned in a similar fashion taking edge effects into account. People travel outside of the 500,000 person population with probability 10^{-3} per day (2).

   Fig. S5: Empirical probability mass function of the number of sufficient contacts (degree distribution) that a person has with others. This distribution was generated from 100 realizations of the weighted person-to-person contact graph.

   Fig. S6: Part of one realization of the contact graph. The left panel shows people as

squares, while the red lines show who they make contact with. The white lines show part of a triad. In the bottom right panel, we see the triangle in more detail. At the top of the triangle is person id#1565 who is a young adult that makes sufficient contact with two other people (red lines). These two other people make sufficient contact with one another (yellow line). The middle aged adult has degree 2 and the clustering coefficient for the triad is 1.

Fig. S7: Part of one realization of the contact graph. The left panel shows people as squares, while the red lines show who they make contact with. The white lines show how one person makes sufficient contact with five other people. In the bottom right panel, we see this in more detail. In the middle is person id#1568 who is a young adult that makes sufficient contact with five other people (red lines). None of these five other people make contact with each other. The middle aged adult has degree 5 and the clustering coefficient for the group is 0.

Fig. S8: Part of one realization of the contact graph. The left panel shows people as squares, while the red lines show who they make contact with. The white lines show how one person makes sufficient contact with 17 other people. In the bottom right panel, we see this in more detail. In the middle is person id#550 who is an elementary school child that makes sufficient contact with 17 other people (red lines). These contacts make a number of sufficient contacts with others in the group (yellow lines). The elementary school child has degree 17 and the clustering coefficient for the group is 0.13.

Fig. S9: Average number of sufficient contacts made per day between an average person in each age group from the entire population of 500,000.

Fig. S10: The basic reproductive number, R_0, of a newly emergent influenza strain. R_0 is defined as the number of secondary infections produced by a random infected person in a fully susceptible population. To calculate R_0, we assumed a scenario in which one randomly chosen, unvaccinated infected person was seeded into a population where everyone else's ability to transmit was 0, and then counted the number of secondary infections. This was repeated 1000 times. The frequency is the number of times out of 1000 that the given number of secondary cases occurred (*e.g.*, there where 0 secondary cases for 292 of the stochastic replications). $R_0 = 1.4$ in this case.

Fig. S11. Sensitivity analysis for the average number of cases per 1,000 people when varying the percentage of TAP started 14 days after the first case at different values of R_0.

Fig. S12: Sensitivity analysis for the average number of cases per 1,000 people when varying the percentage of GTAP started 14 days after the first case at different values of R_0.

Fig. S13: Sensitivity analysis for the average number of cases per 1,000 people when varying the varying the percentage of households and neighborhood clusters in quarantine started 14 days after the first case at different values of R_0. The number of cases per 1000 does not always decrease monotonically with increasing intervention coverage for small differences due to stochastic variability.

Fig. S14: Sensitivity analysis for the average number of cases per 1,000 people when varying the number of days after the first detected case that the intervention process is triggered when $R_0 = 1.4$

Fig. S15: Sensitivity analysis for the average number of cases per 1,000 people when varying the delay for 80% TAP from 1 - 5 days after an index case is detected in close contact mixing groups when $R_0 = 1.4$.

Fig. S16: Sensitivity analysis of the effectiveness of 90% GTAP and 80% TAP started 14 days after the first case at different levels of AVE_S, when $\text{AVE}_I = 0.62$, $\text{AVE}_D = 0.60$ and $R_0 = 1.4$.Also shown is the resulting variation in the $\text{AVE}_{SD} = 1-(1-\text{AVE}_S)(1-\text{AVE}_D)$.

Fig. S17: Sensitivity analysis of the effectiveness of 90% GTAP and 80% TAP started 14 days after the first case at different levels of AVE_D, when $\text{AVE}_I = 0.62$, $\text{AVE}_S = 0.30$ and $R_0 = 1.4$. Also shown is the resulting variation in the $\text{AVE}_{SD} = 1-(1-\text{AVE}_S)(1-\text{AVE}_D)$.

Fig. S18: Sensitivity analysis of the effectiveness of 90% GTAP and 80% TAP started 14 days after the first case at different levels of AVE_I, when $\text{AVE}_S = 0.30$, $\text{AVE}_D = 0.60$ and $R_0 = 1.4$.

Fig. S19: Sensitivity analysis of the effectiveness of 80% TAP with 70% pre-vaccination started 14 days after the first case at different levels of VE_I, when $\text{VE}_S = 0.30$ and $R_0 = 1.4$.

Table S1: Household composition

# Persons	% of Households
1	7.30%
2	15.30%
3	24.10%
4	29.80%
5	17.60%
6	3.00%
7	1.50%
8	0.70%
9	0.30%
10+	0.30%

Table S2: Age distribution

Age (Yrs)	% of Population
<6mos	0.90%
6mos - 4 yrs	8.20%
5-10	11.00%
11-14	5.50%
15-17	5.50%
18-44	44.40%
45-64	18.30%
65+	6.30%

Table S3: Hospitalization probabilities for a flu-like illness

Age	Probability
<1	0.0023
2-4	0.0025
5-17	0.0008
18-64	0.0007
65+	0.0015

Table S4: Model calibration

	Illness Attack Rate		
		Modeled	
	Asian A (H2N2)	New Pandemic	Hong Kong A (H3N2)
	1957-58	Strain	1968-69
Young Children	35%	32%	34%
Older Children	62%	46%	35%
Adults	24%	29%	33%
Overall	33%	33%	34%

Table S5: Daily contact probabilities by mixing group

	Children					
	Pre-School		School			
Contact group	Small Playgroup	Large Daycare	Elementary	Middle	High	Adults
Small playgroups	0.35					
Large playgroups		0.25				
Elementary school			0.062			
Middle school				0.062		
High school					0.06	
Family						
Child	0.60	0.60	0.60	0.60	0.60	0.30
Adult	0.30	0.30	0.30	0.30	0.30	0.40
Neighborhood Cluster						
Child	0.15	0.15	0.15	0.15	0.15	0.08
Adult	0.08	0.08	0.08	0.08	0.08	0.10
Hospital						
Flu ward						
Worker-worker						0.01250
Patient-worker	0.01000	0.01000	0.01000	0.01000	0.01000	0.01000
Patient-visitor	0.01000	0.01000	0.01000	0.01000	0.01000	0.01000
Other wards						0.00250
Workgroup						0.115
Social Groups	0.0024	0.0024	0.00255	0.00255	0.00255	0.0048

Table S6: Average daily contact degree, by mixing group

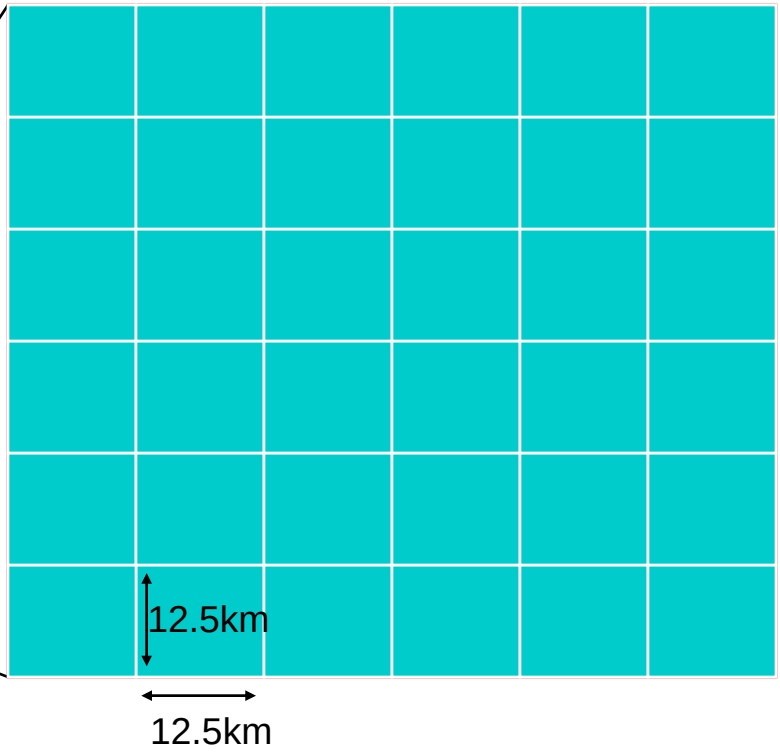
	Children					
	Pre-School		School			
Contact group	Small Playgroup	Large Daycare	Elementary	Middle	High	Adults
Small playgroups	0.98					
Large playgroups		5.2				
Elementary school			7.2			
Middle school				5.9		
High school					4.1	
Family						
Child	0.37	0.37	0.37	0.37	0.37	0.72
Adult	0.48	0.48	0.48	0.48	0.48	0.56
Neighborhood Cluster						
Child	0.56	0.56	0.56	0.56	0.56	0.75
Adult	0.38	0.38	0.38	0.38	0.38	0.84
Hospital						
Flu ward						
Worker-worker						1.9
Patient-worker	-	-	-	-	-	-
Patient-visitor	-	-	-	-	-	-
Other wards						1.4
Workgroup						2.3
Social Groups	0.27	0.27	0.34	0..34	.34	1.14

Table S7: Sources of infection

Sources of Infection	% of Infections
Family	28
Household Cluster	20
Daycare/Playgroup	4
School	21
Workgroup	18
Social Groups	11
Hospital	0.01
Total	102.01*

* Infection could occur in two or more locations on the same day

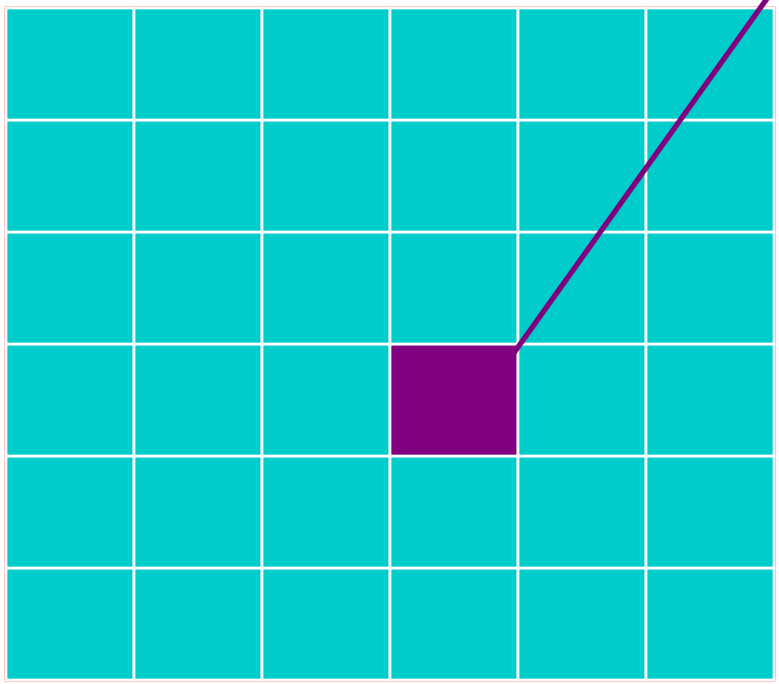
Fig S1



Population characteristics

- 36 localities each of size ~14,000 people
- Total area: 75 km X 75 km = 5,625 km^2
- Population density ~89/km^2

Fig S2



Locality characteristics

- ~ 28 villages, each of size ~ 138 households, ~ 500 people
- Villages are clustered

Within village clusters:

- Household are clustered
- Small & large playgroups
- Elementary, lower-secondary and upper-secondary school mixing groups
- Social groups
- Work groups

Fig S3

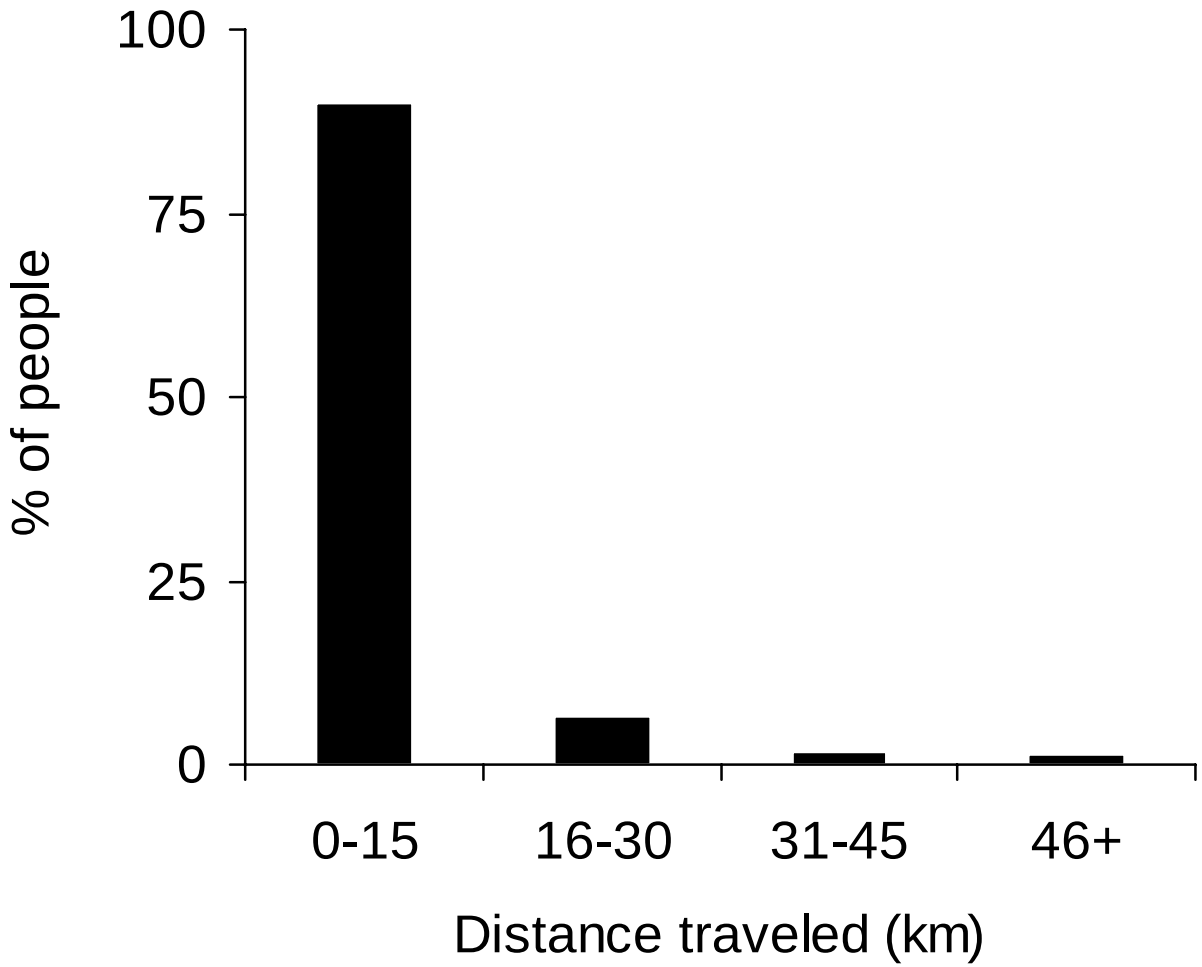


Fig S4

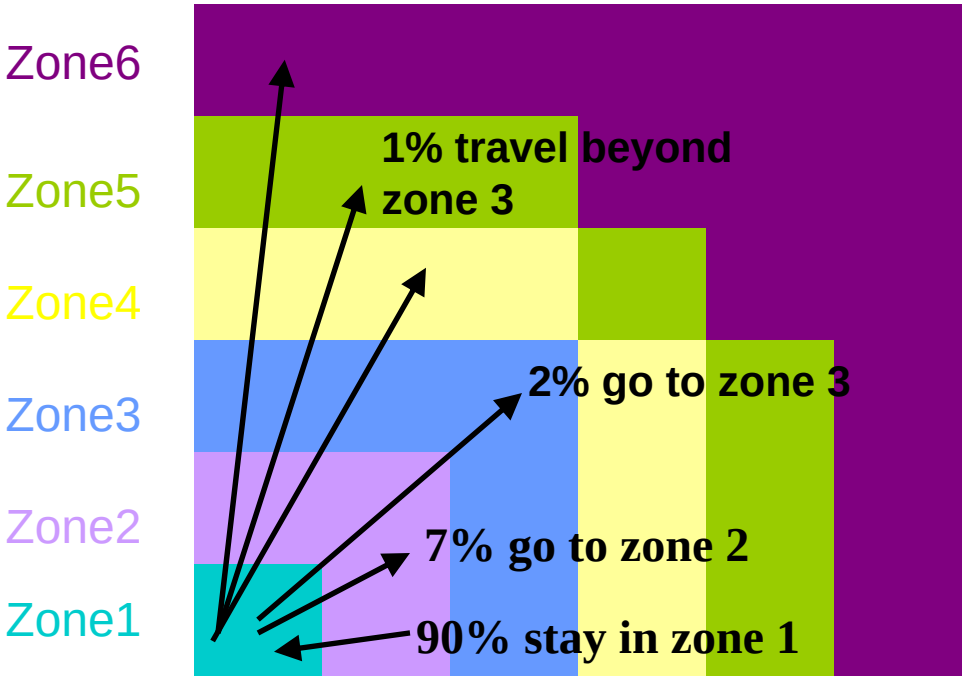
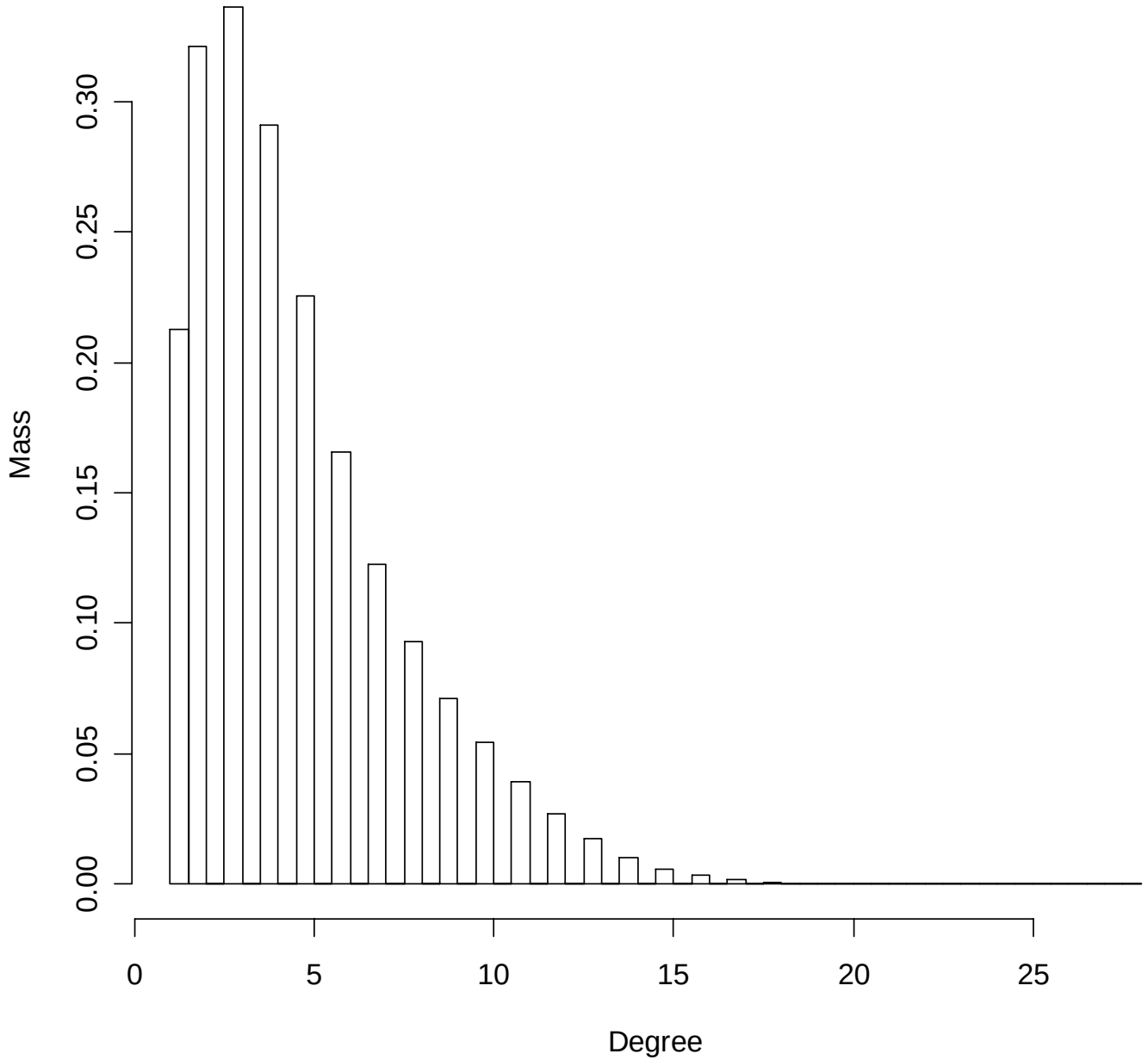

Zone6
Zone5
Zone4
Zone3
Zone2
Zone1
1% travel beyond zone 3
2% go to zone 3
7% go to zone 2
90% stay in zone 1

Fig S5



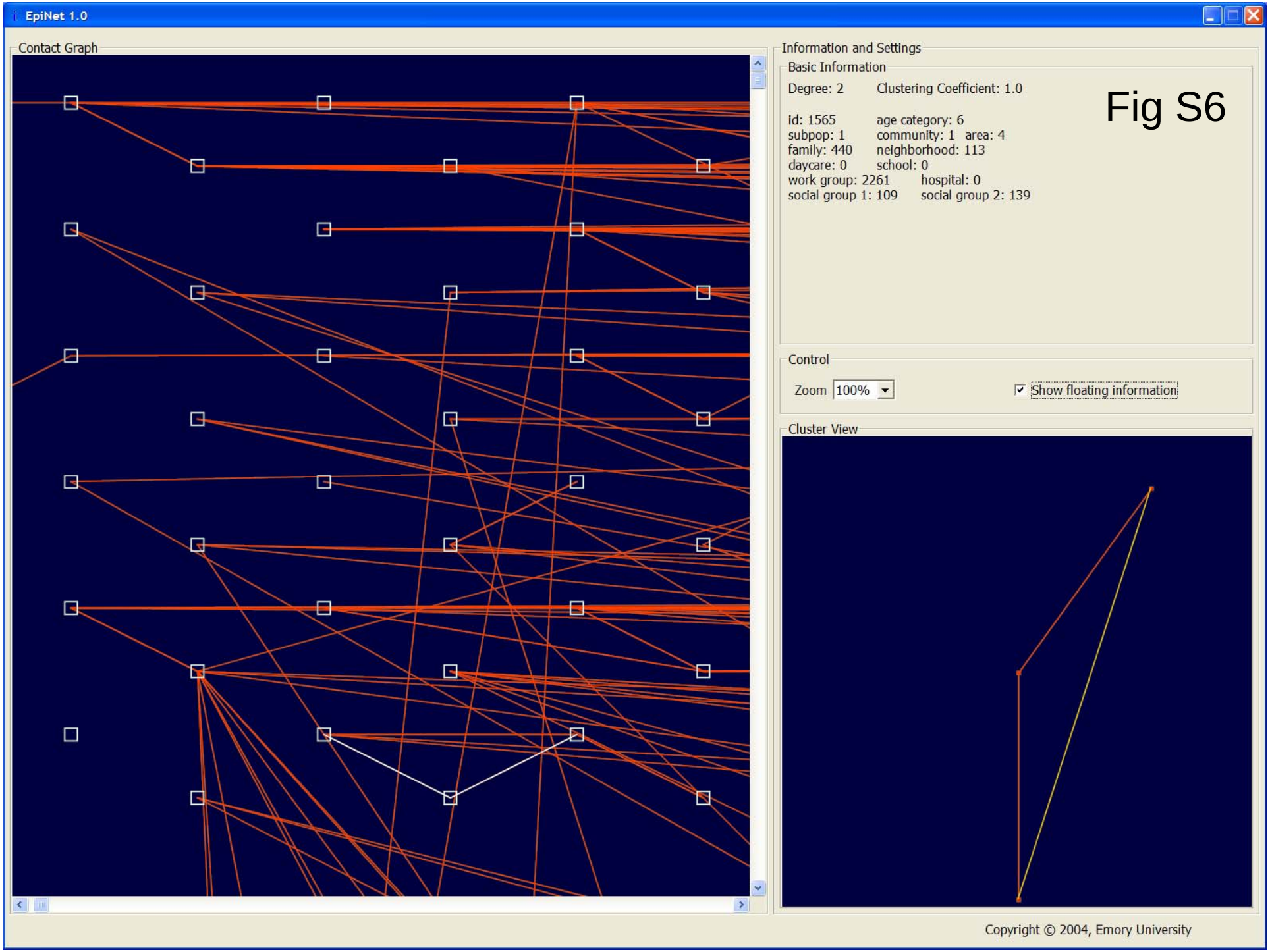

EpiNet 1.0
Contact Graph
Information and Settings
Basic Information
Degree: 2
Clustering Coefficient: 1.0
id: 1565
age category: 6
subpop: 1
community: 1
area: 4
family: 440
neighborhood: 113
daycare: 0
school: 0
work group: 2261
hospital: 0
social group 1: 109
social group 2: 139
Control
Zoom
100%
Show floating information
Cluster View
Copyright © 2004, Emory University

Fig S6

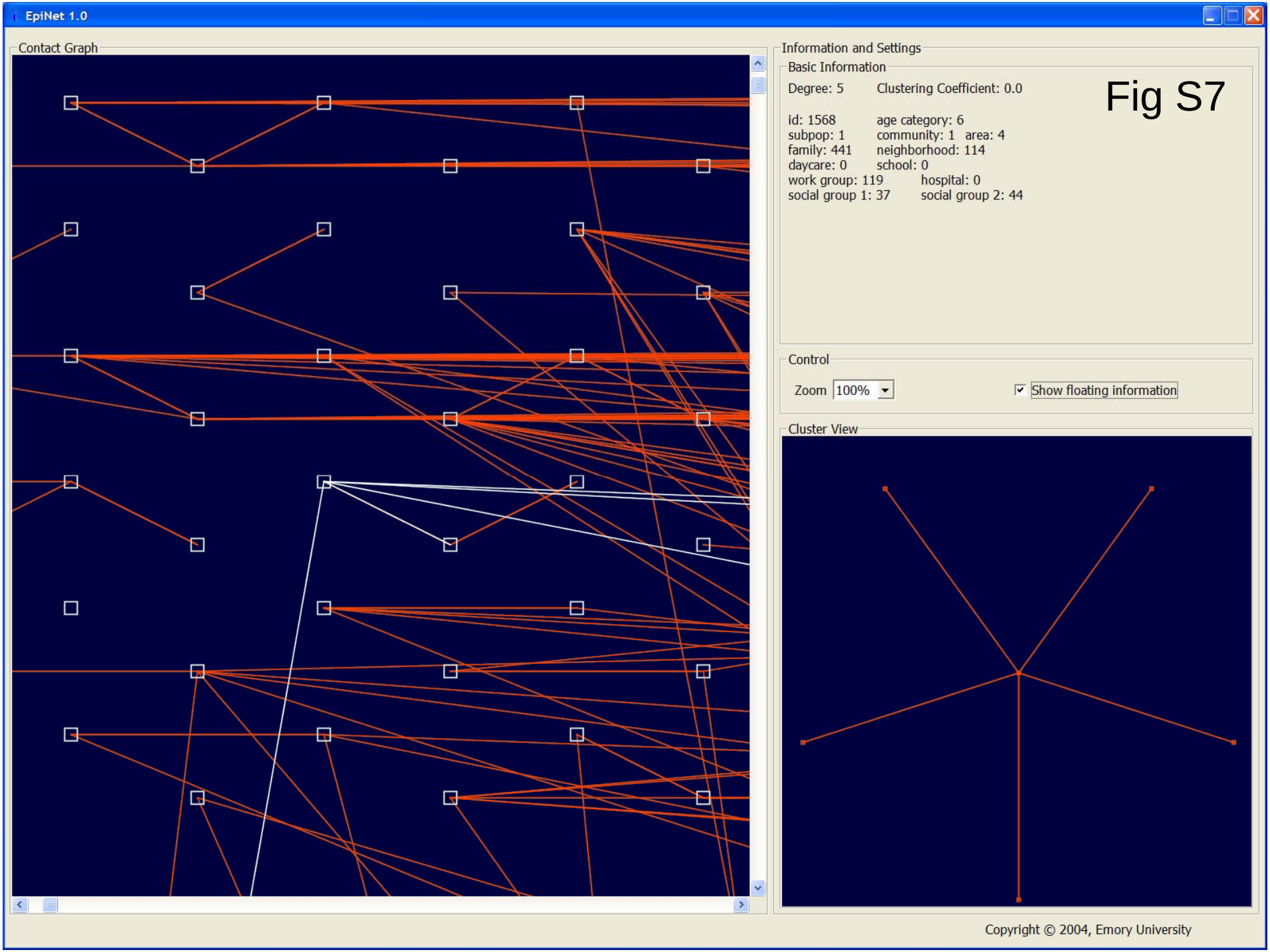

EpiNet 1.0
Contact Graph
Information and Settings
Basic Information
Degree: 5
Clustering Coefficient: 0.0
id: 1568
age category: 6
subpop: 1
community: 1
area: 4
family: 441
neighborhood: 114
daycare: 0
school: 0
work group: 119
hospital: 0
social group 1: 37
social group 2: 44
Control
Zoom
100%
Show floating information
Cluster View
Copyright © 2004, Emory University

Fig S7

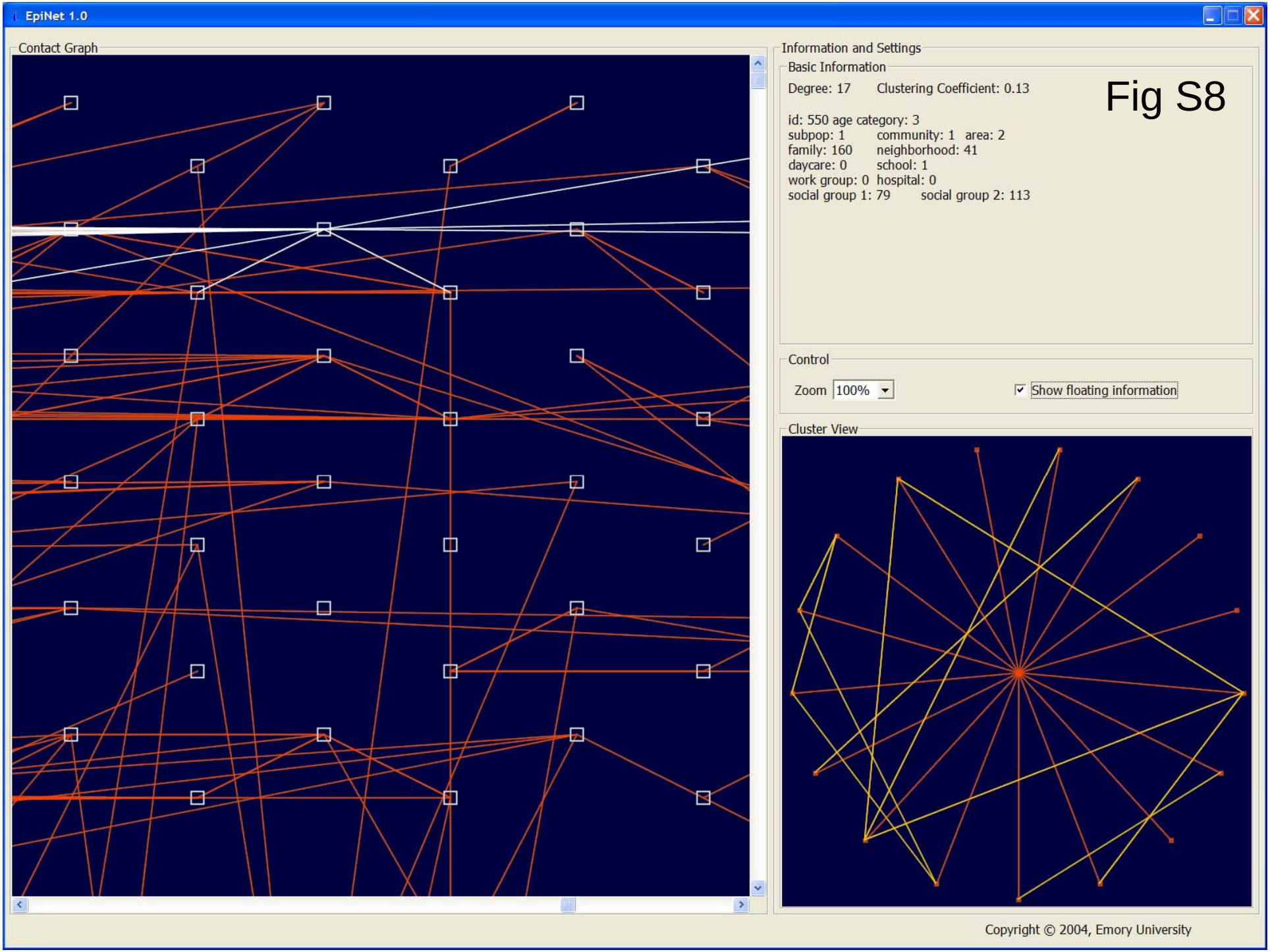

EpiNet 1.0
Contact Graph
Information and Settings
Basic Information
Degree: 17 Clustering Coefficient: 0.13
id: 550 age category: 3
subpop: 1 community: 1 area: 2
family: 160 neighborhood: 41
daycare: 0 school: 1
work group: 0 hospital: 0
social group 1: 79 social group 2: 113
Control
Zoom 100%
Show floating information
Cluster View
Copyright © 2004, Emory University

Fig S8

Fig S9: Average sufficient contact matrix

	From							
age group	< 6 m	6 m - 4	5 – 10	11– 14	15 -17	18-44	45-64	65+
< 6 m	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
6 m - 4	1.0	3.9	1.0	1.0	1.0	1.1	1.1	1.1
5 - 10	1.1	1.1	7.5	1.1	1.1	1.1	1.1	1.1
To 11 - 14	1.0	1.0	1.0	5.9	1.0	1.0	1.0	1.1
15 - 17	1.0	1.0	1.0	1.0	4.2	1.0	1.0	1.1
18 - 44	1.3	1.4	1.4	1.4	1.4	2.2	2.1	2.4
45 - 64	1.1	1.1	1.1	1.1	1.1	1.4	1.4	1.4
65+	1.1	1.1	1.1	1.1	1.1	1.2	1.2	1.2

Fig S10

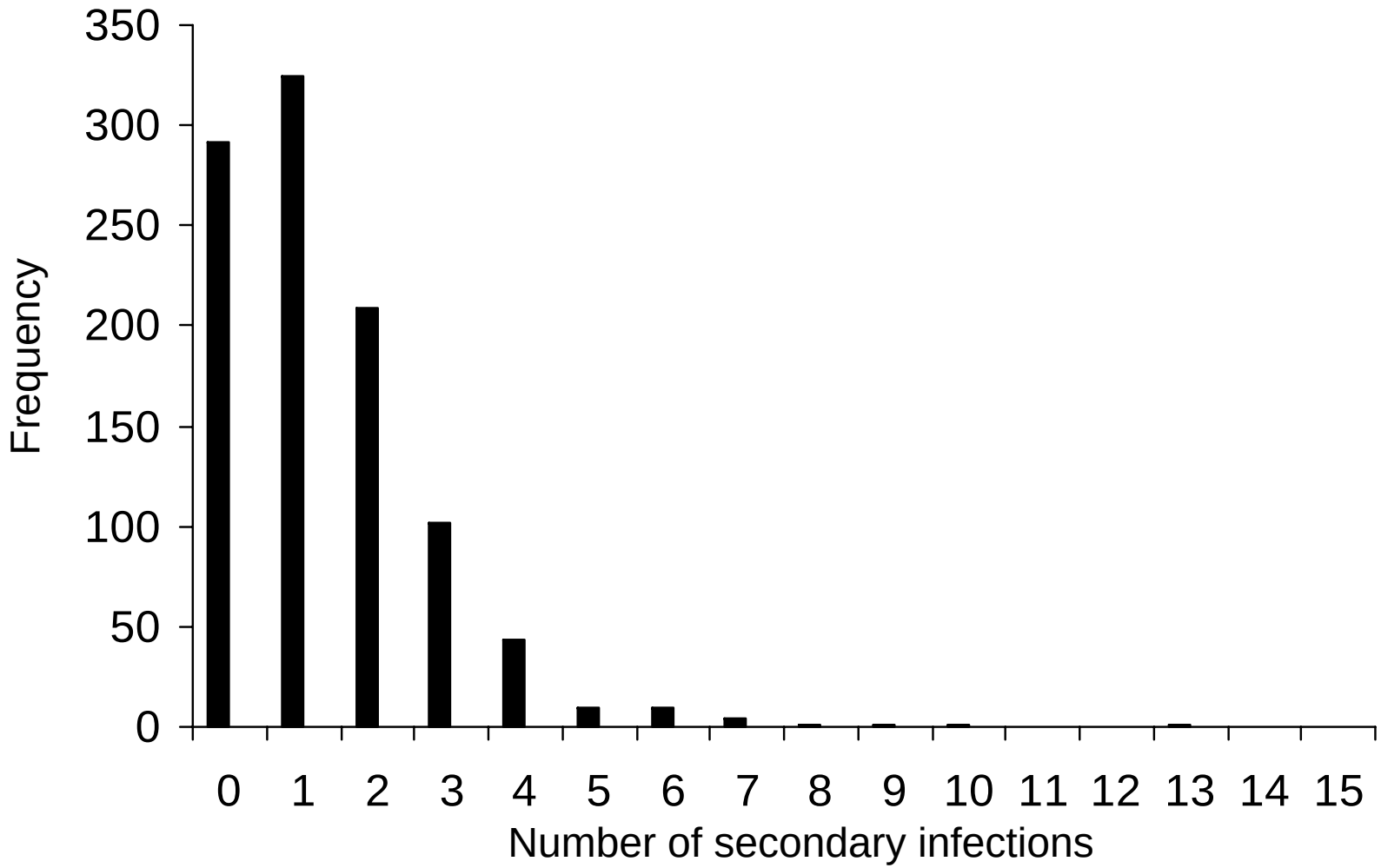

Frequency
350
300
250
200
150
100
50
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Number of secondary infections

Fig. S11

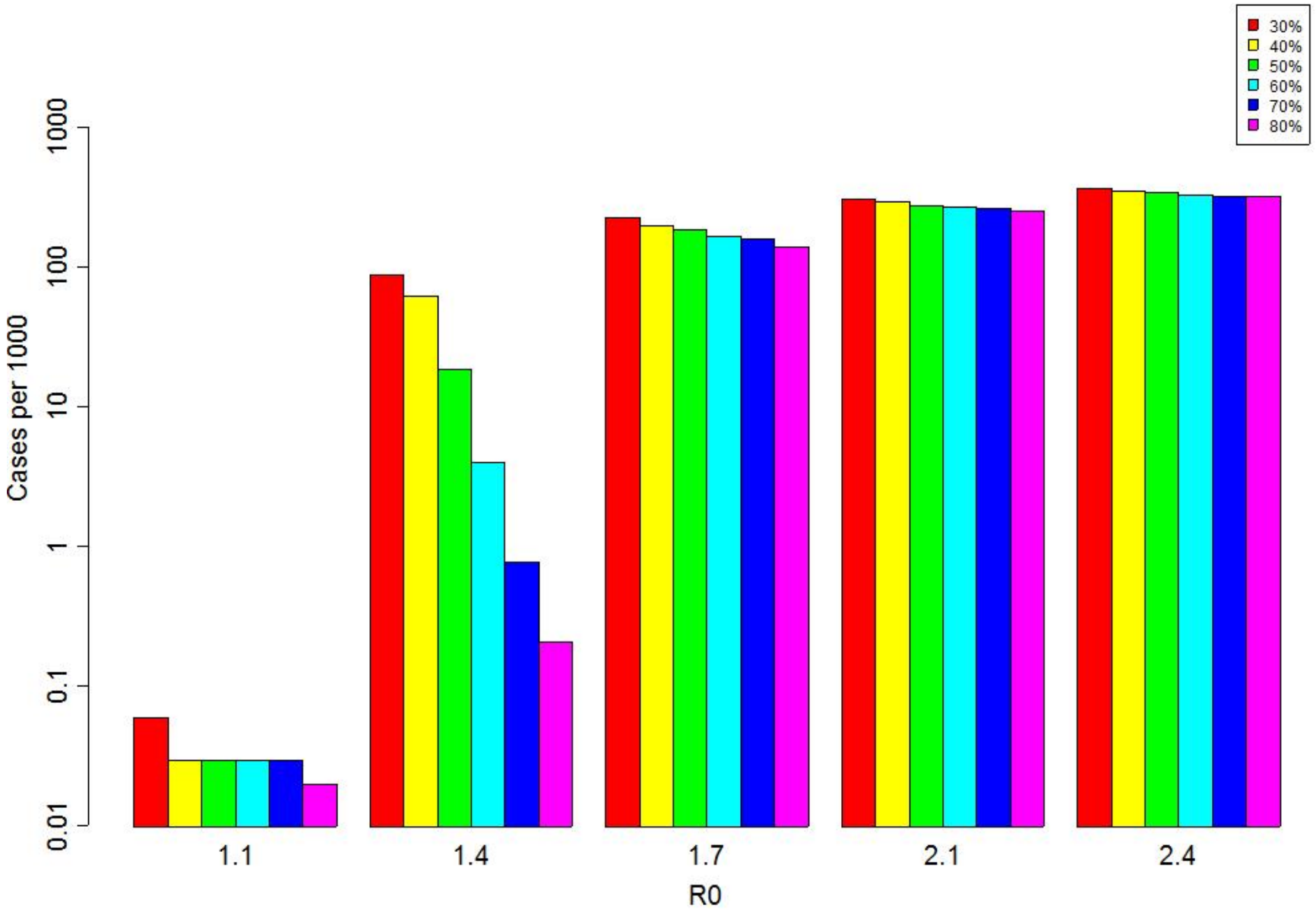

30%
40%
50%
60%
70%
80%
Cases per 1000
0.01
0.1
1
10
100
1000
1.1
1.4
1.7
2.1
2.4
R0

Fig. S12

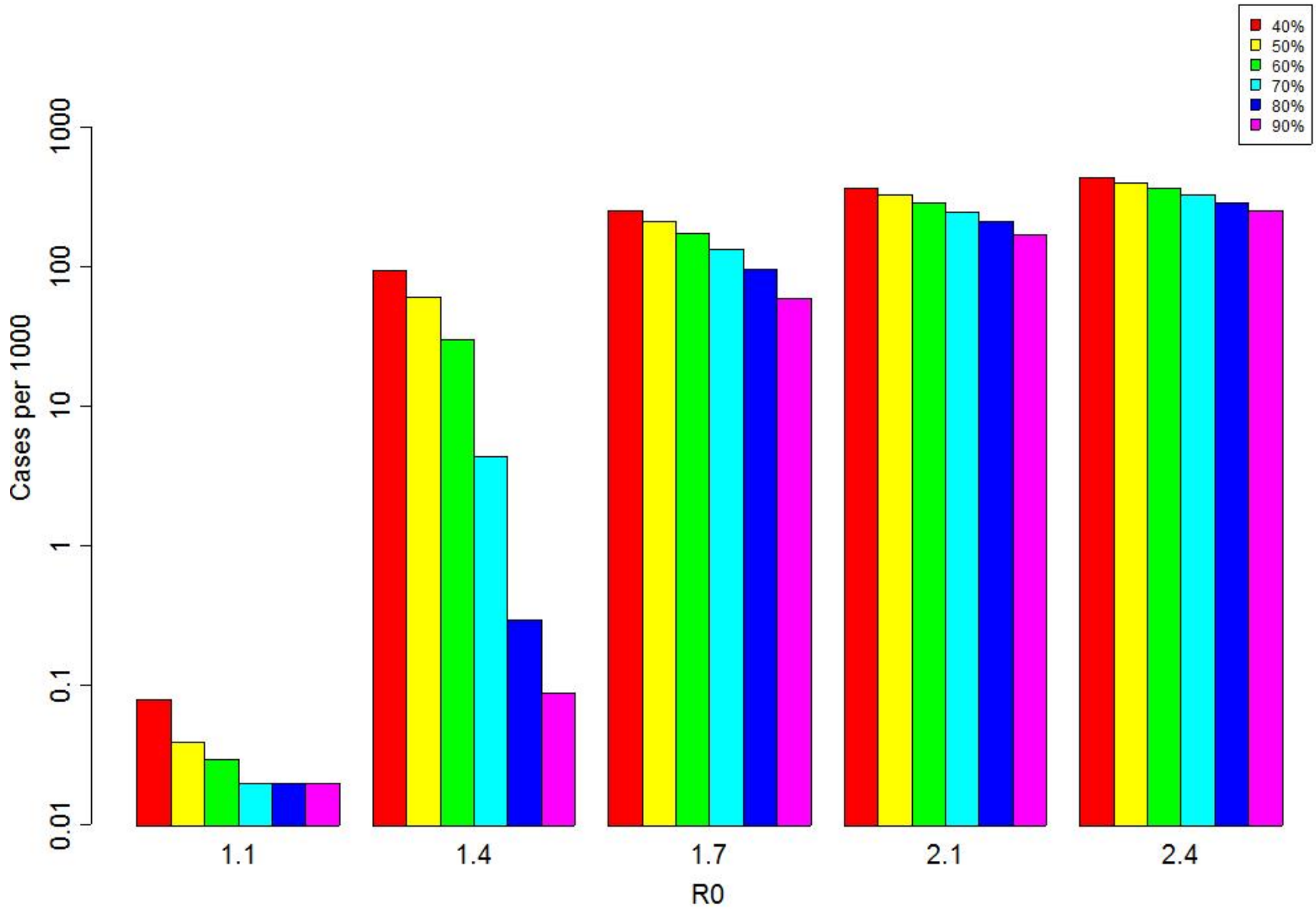

40%
50%
60%
70%
80%
90%
Cases per 1000
0.01
0.1
1
10
100
1000
1.1
1.4
1.7
2.1
2.4
R0

Fig. S13

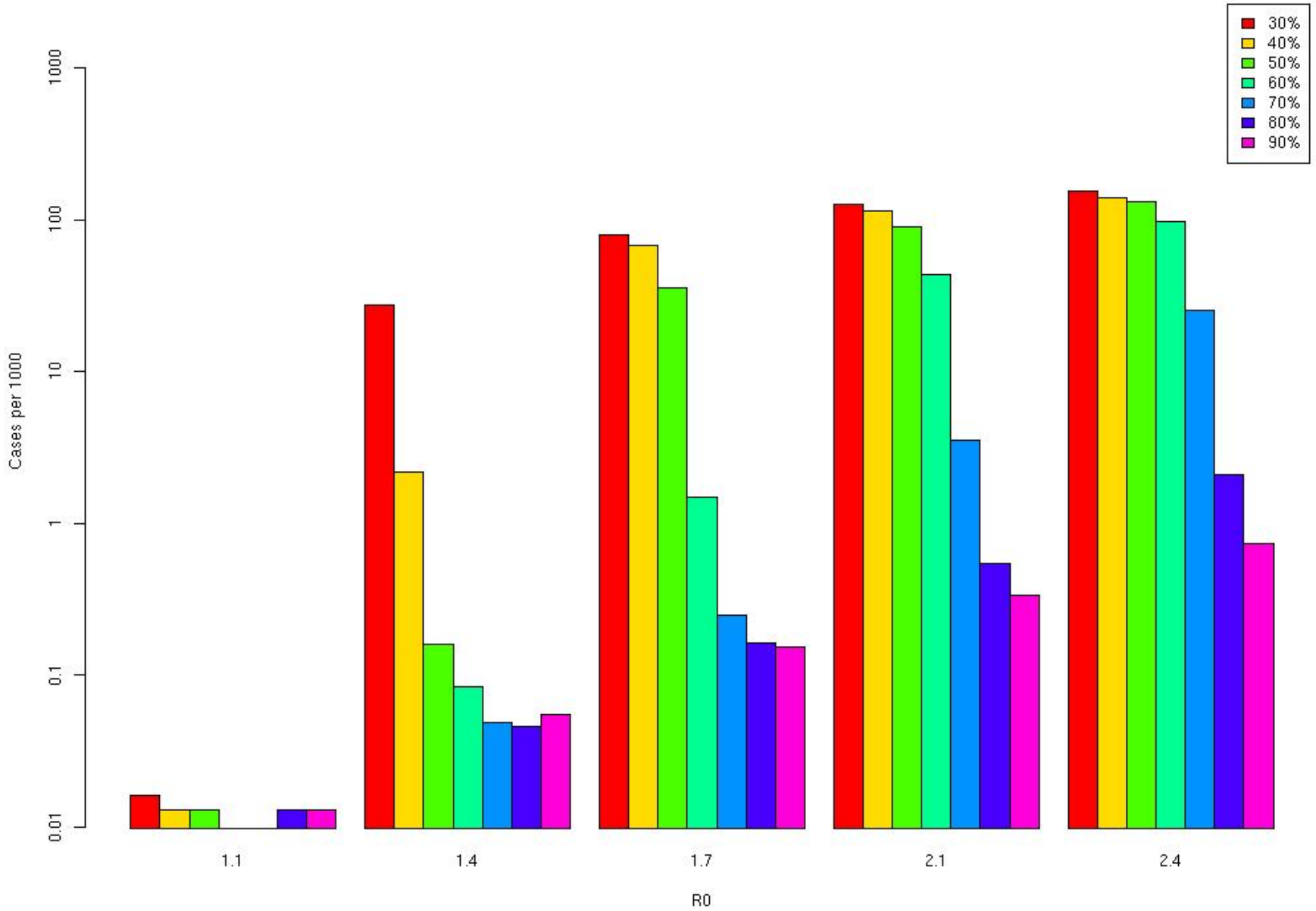

30%
40%
50%
60%
70%
80%
90%
Cases per 1000
0.01
0.1
1
10
100
1000
1.1
1.4
1.7
2.1
2.4
R0

Fig. S14

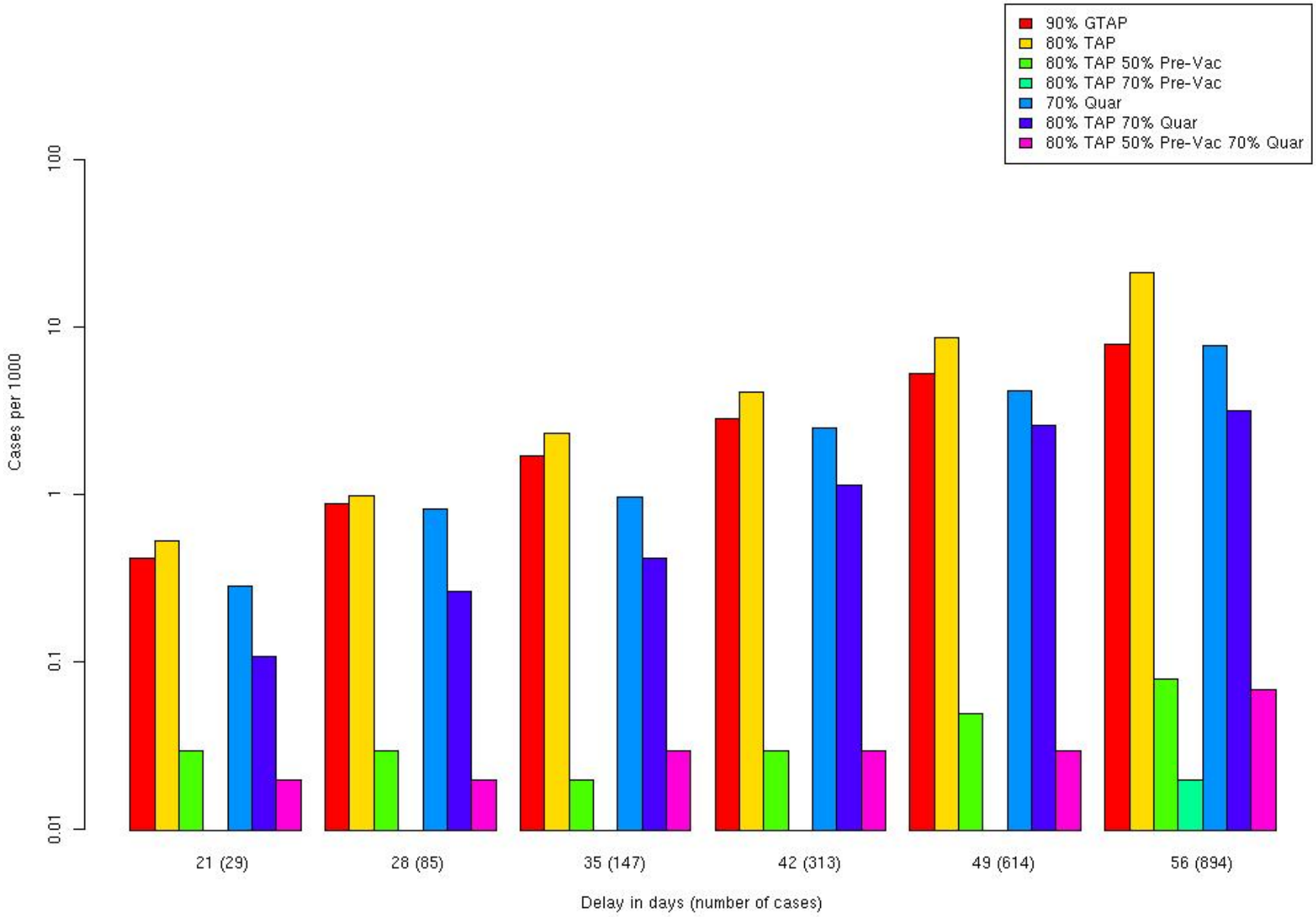

90% GTAP
80% TAP
80% TAP 50% Pre-Vac
80% TAP 70% Pre-Vac
70% Quar
80% TAP 70% Quar
80% TAP 50% Pre-Vac 70% Quar
Cases per 1000
0.01
0.1
1
10
100
21 (29)
28 (85)
35 (147)
42 (313)
49 (614)
56 (894)
Delay in days (number of cases)

Fig. S15

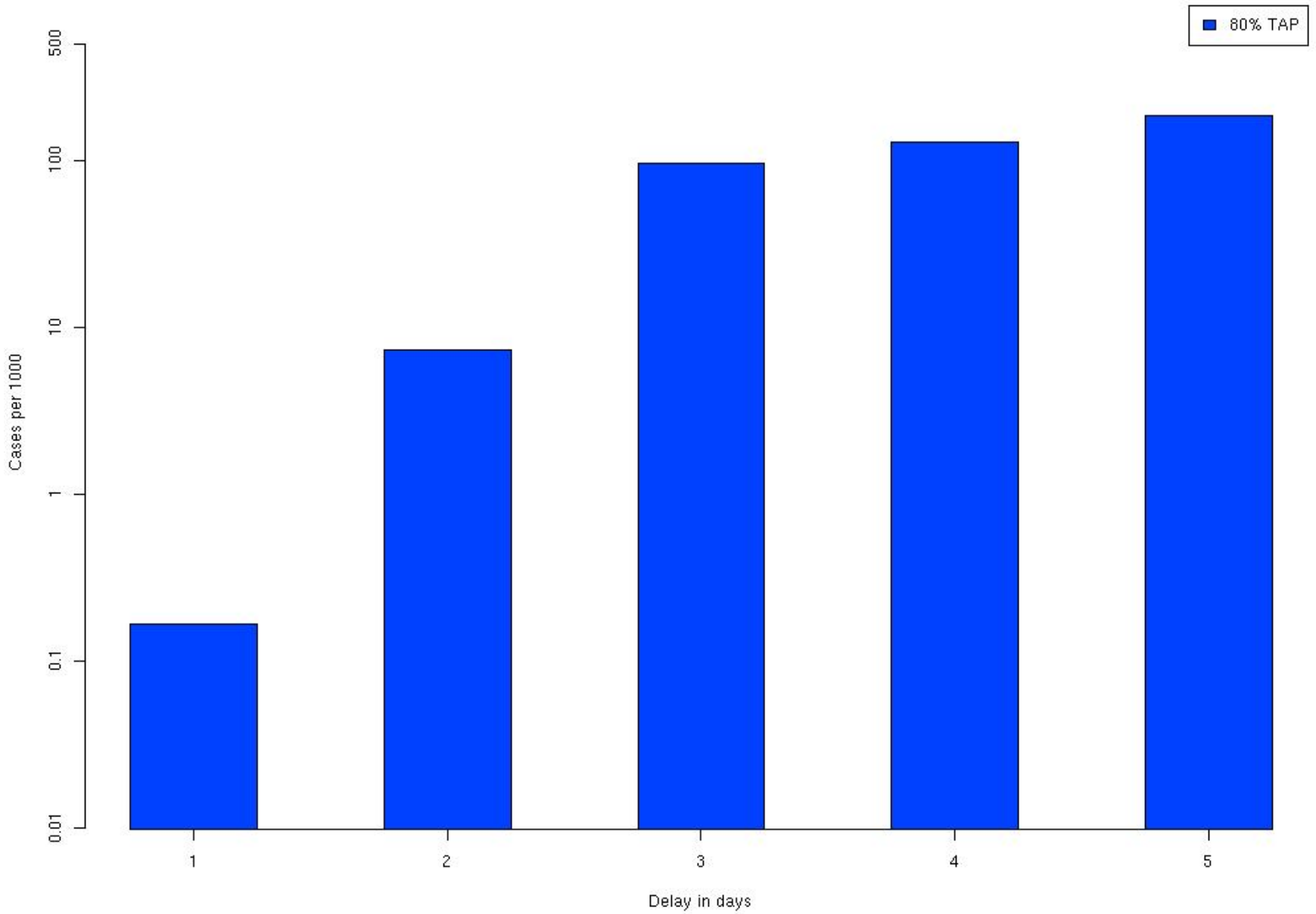


Fig S16

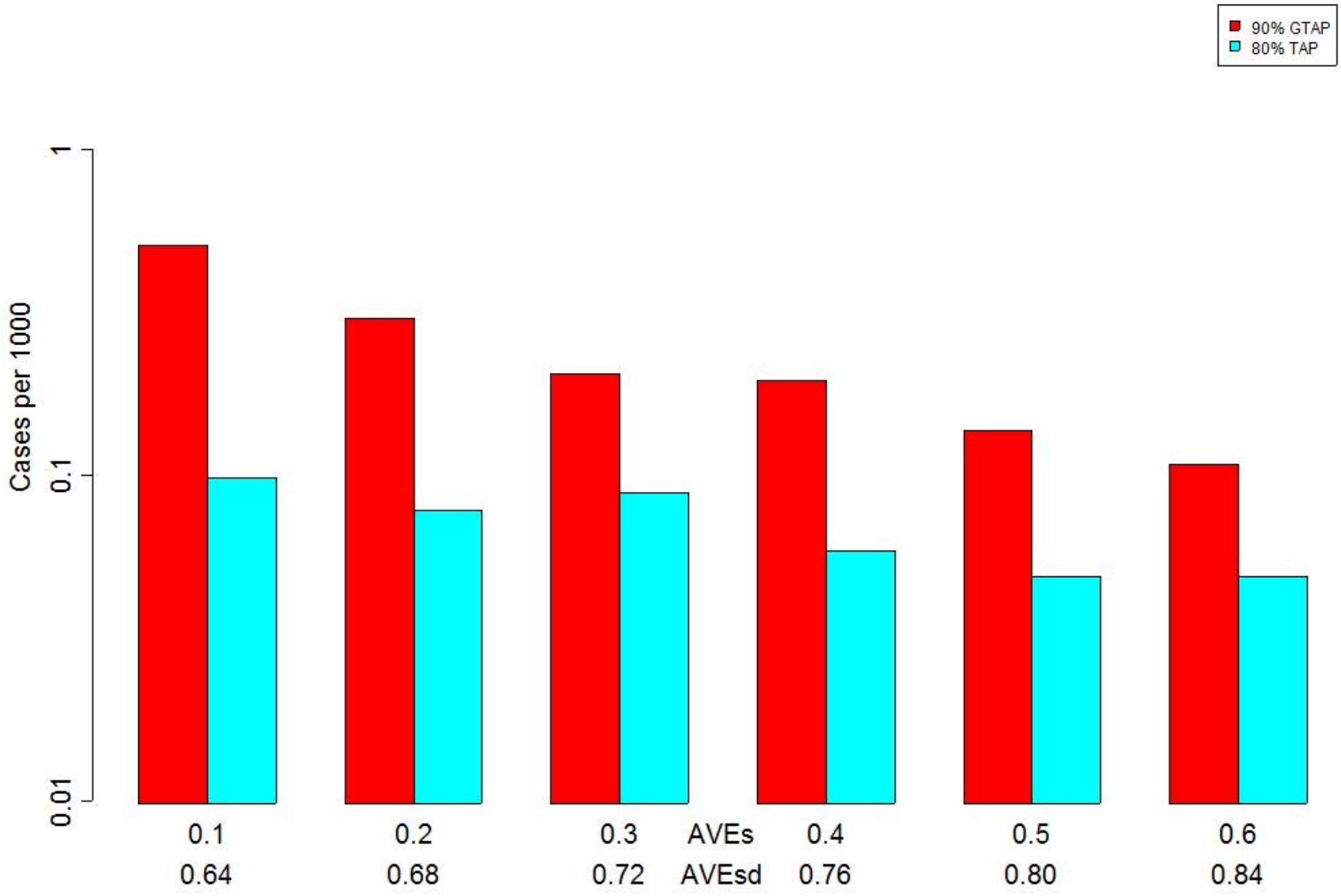

90% GTAP
80% TAP
Cases per 1000
1
0.1
0.01
0.1
0.2
0.3
AVEs
0.4
0.5
0.6
0.64
0.68
0.72
AVEsd
0.76
0.80
0.84

Fig S17

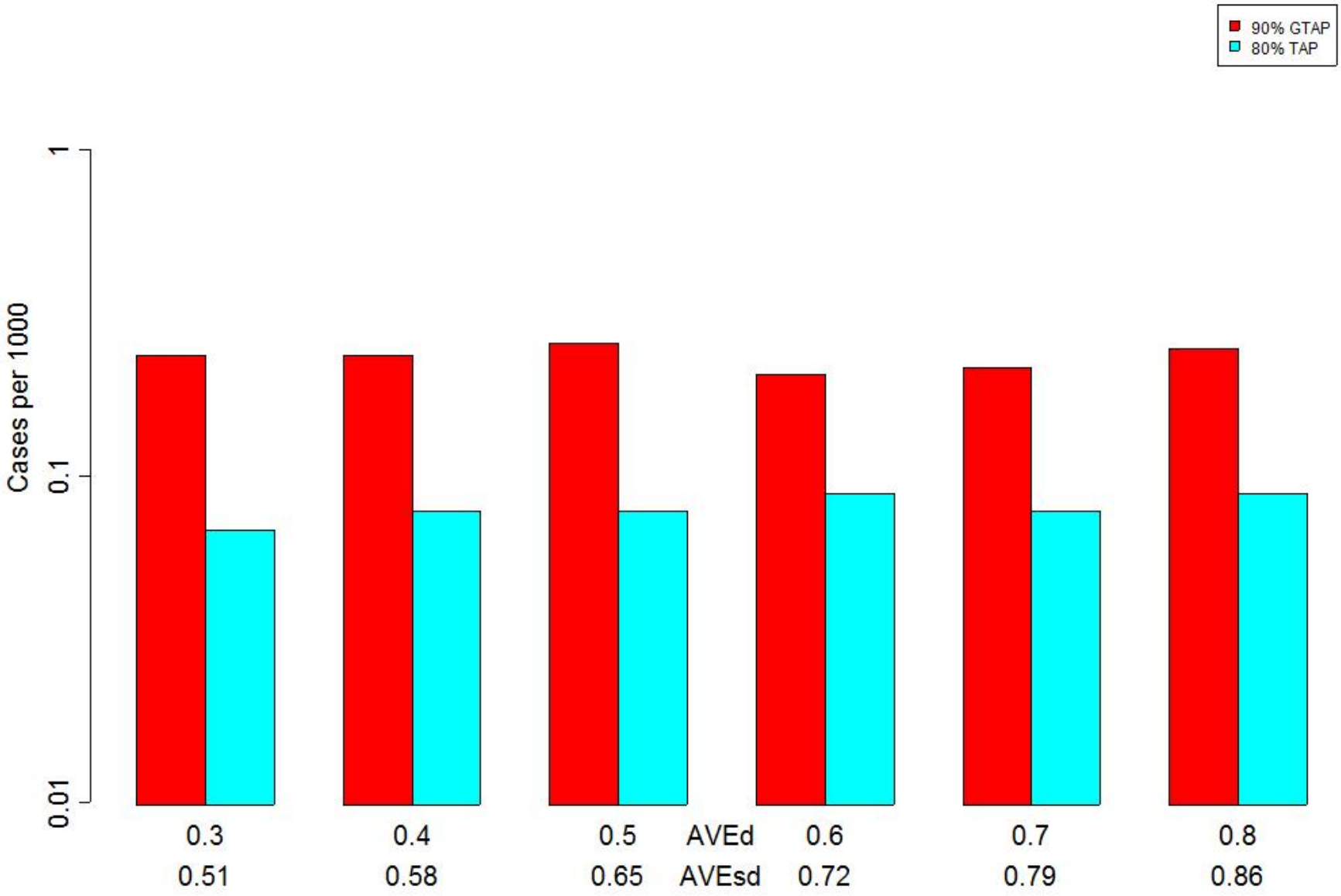

90% GTAP
80% TAP
Cases per 1000
1
0.1
0.01
0.3
0.4
0.5
AVEd
0.6
0.7
0.8
0.51
0.58
0.65
AVEsd
0.72
0.79
0.86

Fig S18

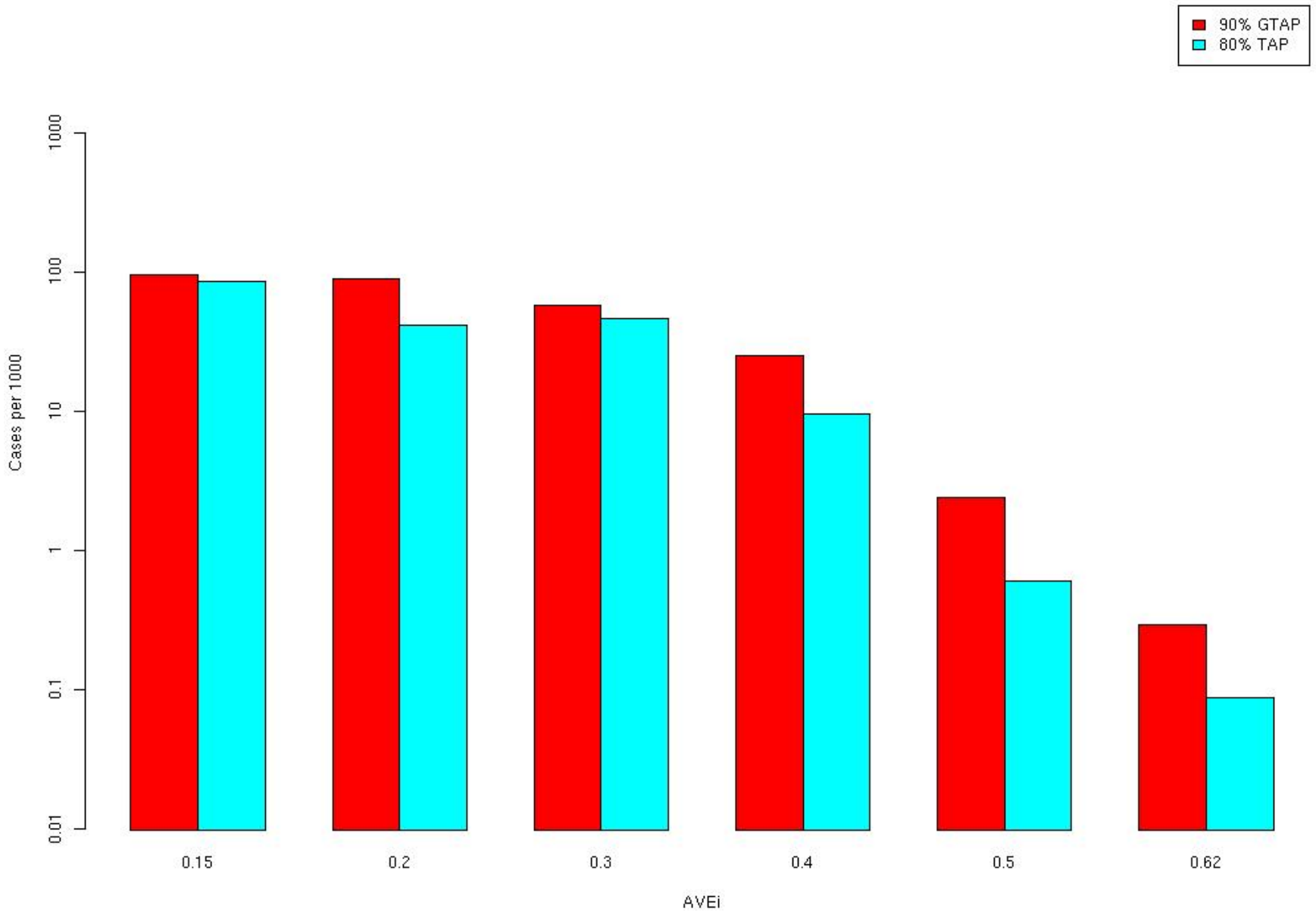
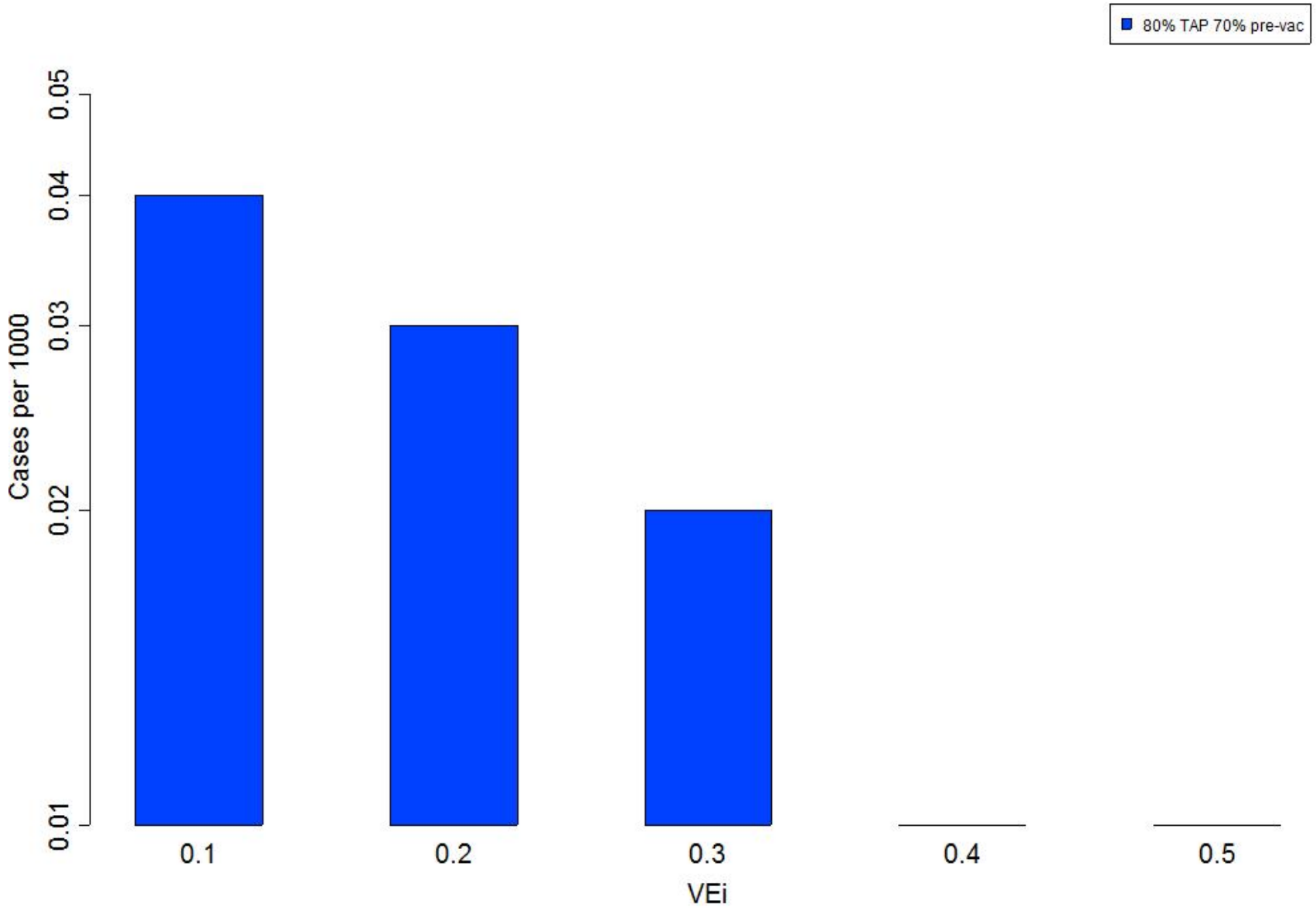

90% GTAP
80% TAP
Cases per 1000
1000
100
10
1
0.1
0.01
0.15
0.2
0.3
0.4
0.5
0.62
AVEi

Fig S19


80% TAP 70% pre-vac
Cases per 1000
0.01
0.02
0.03
0.04
0.05
0.1
0.2
0.3
0.4
0.5
VEi