Quantifying How Social Mixing Patterns Affect Disease Transmission

Sarah E LeGresley[1,2*], Sara Y Del Valle[1], and James M Hyman[1,3]

[1] Los Alamos National Laboratory, Los Alamos, NM 87545
[2] Department of Physics and Astronomy,
University of Kansas, Lawrence, KS 66046
[3] Department of Mathematics,
Tulane University, New Orleans, LA 70118

Abstract. We analyze how disease spreads among different age groups in a stochastic discrete-event simulation of 19 million synthetic individuals representing the residents of southern California. Quantifying the relative importance of different daily activities of a population is crucial for understanding disease transmission and guiding mitigation strategies. We use the Los Alamos Epidemic Simulation System (EpiSimS) to analyze mixing patterns, approximate the probability of transmission based on the duration of the contact, and estimate a matrix that depicts who acquired infection from whom (WAIFW). We provide some of the first quantitative estimates of how infections spread among different age groups based on the mixing patterns and activities at home, school, and work. Our results support the hypothesis that schools are the most likely place for early transmission and that interventions targeting school-aged children are one of the most effective strategies in fighting an epidemic.

Keywords: Infectious diseases; Mathematical models; Mixing patterns; EpiSimS; WAIFW matrix

1 Introduction

The spread of infectious diseases depends upon the contact patterns among people in the population. Mathematical models predicting the spread of a disease must accurately account for the mixing patterns within the population. Once the relationship between the disease spread and the contact structure is understood, the information can be used to identify activities where the disease is most likely to be transmitted and to indicate where interventions might be most effective. The lack of detailed survey data quantifying how people of different ages mix has been one of the limiting factors for accurately modeling disease transmission.

Although there is limited data available on the contact patterns in the real-world [9, 18, 22], there are sophisticated computer simulations that incorporate realistic mixing patterns to match real-world behavior [21]. We analyze the social

* Corresponding author. E-mail address: sarah.legresley@gmail.com(Sarah LeGresley).

mixing and contact patterns in a synthetic population created by the stochastic discrete event model, Los Alamos Epidemic Simulation System (EpiSimS) [7, 11, 21]. We then combine these contact patterns with estimates for the susceptibility and infectiousness of the individuals in the population to better understand the roles of social mixing in disease spread.

The mathematical foundation has been well developed for epidemic models where there is strong biased mixing between different age groups. The non-random mixing formulation include restricted mixing, proportional mixing, preferred mixing, selective mixing, and non-proportionate mixing [3, 14, 15]. These non-random mixing models all require knowledge of the existing mixing patterns in the population. Even though biased mixing epidemic models exist, most common models do not include a detailed account for the mixing between different age groups. One reason is that there is little data to quantify how people of different ages spend time together.

Using age as a metric of mixing is a natural approach since the mixing between ages is highly biased, the course of the disease is often age dependent, and the behavior of the population (e.g., work, school, and play) is directly correlated with age [1]. This paper will provide data for the underlying age-based contact structure that can be directly incorporated into non-random mixing models.

2 Methodology

Following the approaches developed in Del Valle et al. [6], we used EpiSimS, a stochastic discrete-event simulation model, to estimate the number and average duration of daily contacts generated by the population in southern California. EpiSimS is used to simulate movement, activities, and social interactions of individuals based on actual data [7, 11, 21]. The synthetic population for the virtual world is created with the same demographics as the real population as determined by the 2000 U.S. Census Data, including: age, household income, gender, composition of the household, and population density. EpiSimS uses the National Household Transportation Survey (NHTS) to assign activity patterns to the population. The social network, which includes the number of contacts and duration of contacts at each activity, emerges from the simulation.

Disease transmission events can only occur between individuals that occupy the same room at the same time. Nevertheless, each contact has a weight based on the duration of the contact, which in turns modifies the probability of transmission. This detail in disease transmission makes EpiSimS more realistic than macro-scale simulations or other micro-scale simulations where transmission is instantaneous, rather than time-dependent.

The core of this complex epidemic model is the contact structure of the population being modeled. It is through this contact structure that the disease passes from individual to individual and can then be used to predict where the disease is most likely to be transmitted. Also, the structure can be used to define the contact mixing patterns in other disease models that do not have the extensive social contact structured used by EpiSimS.

3 Contact Structure Analysis

We analyzed the population of southern California, which included the counties of Los Angeles, Orange, Riverside, San Bernardino, San Diego, and Ventura. The population consists of about 18.8 million, ranging between 0 and 90 years of age with a median age of 32 and a mean age of 33. A breakdown of the population reveals that preschoolers (ages 0 to 4) are 8.1%, school-aged children (ages 5 to 18) are 22.8%, adults (ages 19 to 65) are 59.8% while seniors (ages 65 and older) are the final 9.2% of the population [21]. The distribution of the ages in Figure 1 reveals a bimodal distribution in the data with the highest peak occurring around the age of 8 and the second, smaller peak, occurring around the age of 35. This bimodal effect may be the results of the increase in population due to the baby boomers and their offspring.

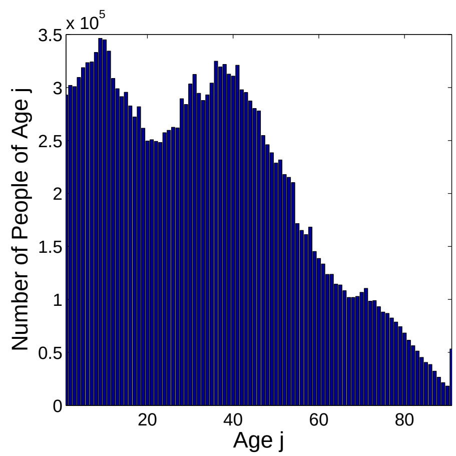

Fig. 1. Age Distribution of the synthetic population for southern California. For the total population of 18,828,569 people, the mean age is 33 while the median age is 32. There are two peaks that occur, one at about 350,000 for 8 year olds and a second at 325,000 for 35 year olds.

3.1 Total Number of Contacts by Activity

We denote the total number of contacts between age groups, matrix C_{ij}, which describes the total number of contacts that a person of age i has with a person of age j. We separated the number of contacts into: 1) children's contacts at school and 2) the rest of the population's contacts, which excludes the contacts between children at school. Figure 2 shows that the aggregated number of contacts between children is on the order of 1,000,000 (top), while the remaining contacts is on the order of 10,000 (bottom). The diamonds along the diagonal on Figure 2 (top) illustrates how children are far more likely to have contacts with children of their same age.

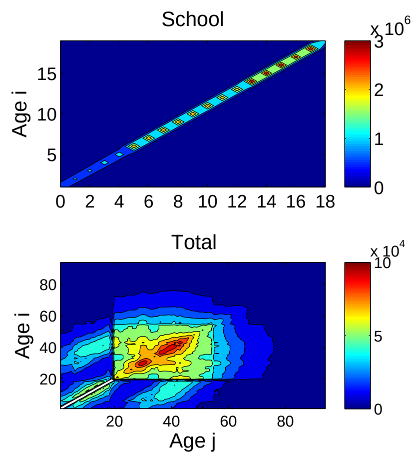

Fig. 2. Total number of contacts between different age groups per day. (**Top**) The number of contacts at schools are the greatest between students of the same age due to the grouping of children into classes by age. The largest number of contacts occurs between teenagers of the same age. (**Bottom**) The number of contacts at all activities but excludes the contacts between children at school. This plot shows that in general, as the age difference increases, the number of contacts decrease. The exception is the weak coupling between children and adults, probably due to the parent-child relationship.

Figure 2 (top) shows that contacts at school occur most frequently between children of the same age. This is due to the stratification, or grouping, of the children into classes at school by age group.

Figure 2 (bottom) shows the number of contacts outside of school. Since children's contacts along the diagonal dominate this plot, we removed them to appreciate the dynamics outside the diagonal. The plot shows that children have the most contacts with children of similar age and fewer contacts as the difference between their ages increases. The contacts between middle age adults, ages $20-60$, have a block pattern in that adults tend to have lots of contacts with adults, with most occurring between adults of the same age. As with contacts between children, adults tend to have more contacts with adults of the same age and fewer contacts as the age difference increases. This assortive (like with like) mixing pattern has been seen by Beutels et al. [5], Del Valle et al. [6], Edmunds et al. [9], Glasser et al. [12], Hens et al. [13], Mossong et al. [18], Newman and Girvan [19], and Wallinga et al. [22]. An exception to this is the weak coupling, or larger number of contacts, that occurs between adults and children, probably due to parent-child relationships. This pattern of strong diagonal and weak coupling is consistent with previous studies including Del Valle et al. [6], Glasser et al. [12], Hen et al. [13], and Mossong et al. [18].

3.2 Average Duration of Contacts by Activity

We denote matrix T_{ij}, the average duration of contact per day in hours that a person of age i had with a person of age j. This matrix is estimated by dividing the total duration of all contacts by the total number of contacts, matrix C_{ij}. Figure 3 confirms that children have the longest contacts with other children their own age, while adults have, on average, shorter contacts over a much broader age range.

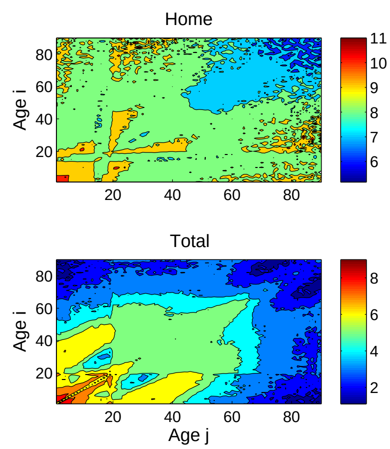

Fig. 3. Average duration of contacts per day in hours. (**Top**) The average duration of contacts at home shows a large variation with age. The average duration of contacts between preschoolers (ages $0 - 4$) is about 10 hours per day while between older adults (ages $80 - 90$) it is only about 5.5 hours per day. (**Bottom**) The average duration of contacts at all activities shows the longest contact durations occurs between children (ages 18 and under) with as much as 9 hours per day, followed by the duration between children and adults (ages $20 - 60$) at around 6 hours, and finally the duration between adults (ages $20 - 60$) at about 5 hours per day.

Figure 3 (top) shows that the average duration of contacts at home vary widely with age. Contacts between preschoolers (ages 0 - 4) are the longest with the average duration being around 10 hours. The shortest average contact durations occur between 80 - 90 year olds with an average of 5.5 hours. This may be due to the fact that more older people live alone.

Figure 3 (bottom) shows the average duration of contacts at all the activities combined. Note that the average contact duration between children (ages 18 and under) are the longest. As seen in the total number of contacts between adults, we see a block for adults (ages 20 - 60) with an average contact duration of around 5 hours. This is probably from contacts between people at work or from spouses in the same household. We observe a weak coupling between children and adults with an average contact duration of around 6 hours, probably due to the parent-child relationship.

3.3 Probability of Transmission by Activity

The probability of transmission, matrix P_{ij}, is based on the duration of contacts between a susceptible in age group i and an infected in age group j. This paper uses the same approach as in [6] where $P_{ij} = 1 - e^{-\sigma T_{ij}}$ and σ is the mean number of transmission events per hour of contact between fully susceptible and fully infectious people. To allow direct comparison with [6], $\sigma = 0.2$ is used. Since this is a Poisson probability distribution with parameter σt, the longer the contact, the greater the probability of transmission.

In the top plot of Figure 4, we show the probability of transmission at home. The probability of transmission between preschoolers (ages 0 - 4) is the highest at around 0.9 and is due to the long duration of their contacts. Between adults, the probability of transmission tends to decrease with increasing age.

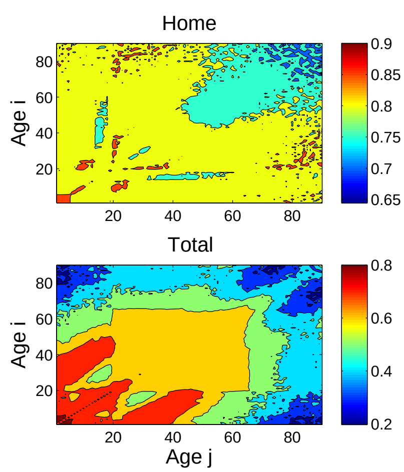

Fig. 4. Probability of Transmission (**Top**) At home, the highest probability of transmission occurs between preschoolers (ages 0 - 4) at 0.9 and is lowest between seniors (ages 80 - 90) at around 0.65. (**Bottom**) At all activities, the highest probability of transmission occurs between preschoolers (ages 0 - 4) at around 0.8. The probability of transmission between school-aged children (ages 5 - 18) is about 0.7 as is the probability of transmission between adults (ages 20 - 50) and children, probably due to the parent-child relationship. The probability of transmission between adults (ages 20 - 65) is about 0.6.

In the bottom plot of Figure 4, we show the probability of transmission at all activities, which reveals two blocks and a weak coupling. For the weak coupling between children and adults, there is a probability of about 0.7. The block for adults (ages 20 - 65) has a probability of transmission of around 0.6. For the block with children, transmissions between school-aged children have a probability of 0.7, while preschoolers have a probability of about 0.8. This is consistent with researchers including Longini et al. [16] and Mikolajczyk et al. [17], which concluded that vaccination of children is an effective mitigation technique in controlling the spread of an infection.

3.4 Who Acquired Infection From Whom by Activity

The transmission matrix, also known as the who acquired infection from whom (WAIFW) matrix, represents the rate β_{ij} at which a susceptible person from age i will be infected by an infectious person from age j. The formula for calculating β_{ij} is $\gamma_{ij} \times \alpha_i \times \xi_{jk} \times P_{ij}$, where γ_{ij} is the average number of contacts per day (which can be calculated by dividing C_{ij} over N_i, where N_i is the total population), α_i is the susceptibility, ξ_{jk} is the infectivity, and P_{ij} is the probability of transmission matrix (see Probability of Transmission by Activity section). In order to simplify the calculations and in keeping with [6], we will assume that α_i and ξ_{jk} are both 1.

Figure 5 (top left) shows the transmission rates at home. The highest transmission rate occurs between children of differing ages. This is probably due to transmission between siblings, who tend to be of different ages. There is also a high transmission rate among adults of a similar age, probably due to spouses of a similar age. Finally, there is a high transmission rate between children and adults.

Figure 5 (top right) shows the transmission rates at school. Transmission rates among children of the same age are by far the largest. This is due to the stratification in EpiSimS placing children of the same age in the same classroom at school. The implication of this finding is that the largest transmission rates occur between teenagers.

Figure 5 (bottom left) illustrates the transmission rates for shopping. This plot shows highly non-symmetrical transmissions. The largest transmission rate is from middle age adults (ages 20 - 60) to older adults (ages 70 - 90). This is probably because older adults have fewer contacts at home and are more likely to be exposed to disease while shopping.

Figure 5 (bottom right) shows the transmission rates at work. Adults have the highest transmission rate at work. This high rate for adults (ages 20 - 60) is expected since most workers are in this age range.

Figure 6 shows a block between adults and the weak coupling between children and adults. In this plot, the transmission rates between children have been removed because the transmission rate between children is dominated by school (see Figure 5) and are significantly larger than transmission rates between any other group. For the adults, the transmission rate of about 0.2 is highest among adults of the same age and decreases with increasing age differences. The exception is where the weak coupling occurs and there is a transmission rate of about 0.1 both from adults to children and from children to adults.

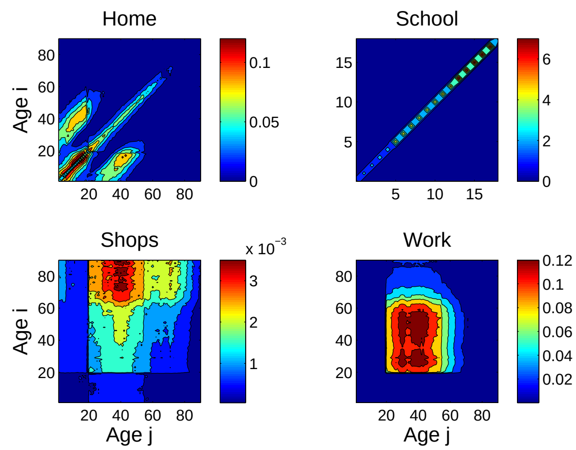

Fig. 5. Transmission Matrix (WAIFW) (**Top left**) At home, the highest transmission rates are between children of different ages, probably siblings, followed by transmission between adults and children, and finally between adults of similar ages. (**Top right**) At school, the largest transmission rates are among students of the same age, with teenagers being the largest. Notice these rates are significantly higher than at home, work, and shops. (**Bottom left**) At shops, the highest transmission rates are from middle age adults (20 - 60) to older adults (ages 70 - 90) though this rate is much lower than the transmission rates at home or work. (**Bottom right**) At work, the highest transmission rates are between adults (ages 20 - 60) which should be expected because they compromise the majority of the work force. Even the highest transmission rates at work are lower than the rates at school but still higher than at shops.

Table 1 shows the aggregated daily transmission rates for the following ages groups: 0 - 4, 5 - 12, 13 - 19, 20 - 29, 30 - 39, 40 - 49, 50 - 59, 60 - 69 and 70 - 90. This is an aggregation of the β_{ij} transmission matrix in Figure 6. For the aggregated transmission rates for home and school see Table 2 and Table 3, respectively.

3.5 Infections by Activity

Table 4 shows the probability of being infected at different activities. Infected children were most likely infected at school, followed by home, and then social recreation, or shopping. Infected adults are most likely to have become infected at home, followed by work and then social recreation, or shopping. Edmunds et al. [10] speculated that the risk of infection is probably greater at home than

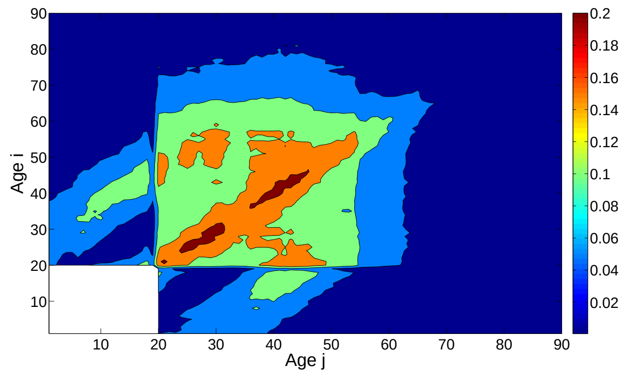

Fig. 6. Transmission Matrix (WAIFW) Total. The transmission rates of children at school dominate, so they have been removed from this plot. The transmission rates are high between children and adults (both from children to adults and adults to children) though the highest rates are between adults (ages 20 - 55) of a similar age.

Table 1. Transmission matrix (WAIFW) of the daily number of adequate contacts per person between the aggregated age groups at all activities. The highest transmission rates are between teenagers (ages 13 - 19).

Age	**0-4**	**5-12**	**13-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-90**
0-4	0.602	0.083	0.042	0.057	0.069	0.030	0.013	0.006	0.002
5-12	0.077	0.744	0.072	0.046	0.091	0.059	0.019	0.009	0.003
13-19	0.046	0.083	0.913	0.057	0.082	0.107	0.041	0.015	0.005
20-29	0.064	0.055	0.058	0.176	0.151	0.146	0.099	0.039	0.011
30-39	0.069	0.096	0.072	0.131	0.173	0.135	0.087	0.039	0.011
40-49	0.033	0.068	0.106	0.143	0.153	0.174	0.098	0.040	0.013
50-59	0.021	0.034	0.059	0.143	0.146	0.143	0.123	0.045	0.013
60-69	0.017	0.026	0.035	0.091	0.103	0.095	0.070	0.050	0.013
70-90	0.010	0.015	0.021	0.042	0.048	0.049	0.036	0.022	0.013

Table 2. Transmission matrix (WAIFW) of the daily number of adequate contacts per person between the aggregated age groups at home. The highest transmission rate is between children (ages 5 - 12).

Age	**0-4**	**5-12**	**13-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-90**
0-4	0.063	0.072	0.028	0.045	0.056	0.018	0.006	0.003	0.001
5-12	0.067	0.086	0.050	0.031	0.073	0.043	0.010	0.004	0.000
13-19	0.030	0.057	0.067	0.023	0.041	0.063	0.017	0.004	0.002
20-29	0.051	0.038	0.024	0.056	0.019	0.021	0.016	0.004	0.001
30-39	0.056	0.078	0.037	0.017	0.043	0.014	0.006	0.005	0.001
40-49	0.020	0.049	0.064	0.021	0.015	0.042	0.011	0.004	0.002
50-59	0.009	0.017	0.024	0.022	0.010	0.014	0.036	0.008	0.002
60-69	0.007	0.011	0.010	0.010	0.012	0.009	0.012	0.022	0.004
70-90	0.003	0.000	0.006	0.004	0.005	0.008	0.007	0.006	0.006

Table 3. Transmission matrix (WAIFW) of the daily number of adequate contacts per person between the aggregated age groups at school. The highest transmission rate is between children (ages 5 - 12). Notice the transmission rates at school are much higher than at home for children.

Age	**0-4**	**5-12**	**13-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-90**
0-4	0.531	0.000	0.000	0.002	0.002	0.002	0.001	0.001	0.000
5-12	0.000	0.641	0.001	0.004	0.005	0.005	0.003	0.001	0.000
13-19	0.000	0.000	0.618	0.004	0.005	0.005	0.003	0.001	0.000

at work. We would expect to see a smaller number of senior adults becoming affected at work but this may be a result of using the household transportation data along with a biased smaller data set towards the working population.

Table 4. If infected, where were you most likely infected? Children ages 19 and under are most likely to have become infected at school followed by home. Adults were most likely to have become infected at home followed by work.

Age	**Home**	**School/Work**	**Social Recreation / Shop**
0 - 4	37.95%	48.19%	13.85%
5 - 12	37.69%	48.95%	13.36%
13 - 19	38.73%	46.73%	14.53%
20 - 29	46.56%	36.00%	17.44%
30 - 39	46.40%	36.55%	17.05%
40 - 49	46.09%	36.59%	17.32%
50 - 59	45.60%	36.50%	17.90%
60 - 69	45.48%	35.83%	18.68%
70 - 90	45.62%	35.18%	19.21%

4 Summary, Discussion, and Conclusions

Using data from EpiSimS, we found the average number of contacts per day followed by the probability of transmission based on the duration of the contact. From the probability of transmission data, the WAIFW matrix was calculated. The WAIFW matrices may be used in deterministic models that stratify the transmission rates by age. We estimated these matrices for all activities combined as well as broken down by activity (home, school, shopping, and work) for southern California. When analyzing all activities combined, we see two blocks occurring in the matrices, one between adults and one between children, as well as the weak coupling between children and adults, probably due to the parent-child relationship.

Finally, we were able to show which activities are more likely to generate secondary infections. For adults, the activity with the highest probability is home followed by work. For children, the activity with the highest probability is school followed by home. Therefore, mitigation techniques targeting children at schools could help halt the spread of disease [8]. This is consistent with researchers having found that mass vaccination would not be necessary [4]. Researchers have also found that vaccinating 80% of children is almost as effective as vaccinating 80% of the population [16].

If the models predictions are used to guide public health policy, models should account for the contact patterns of a population and consider the impact of behavioral changes. Our goal has been to provide estimates for the contact patterns of a synthetic population. Our hope is that these patterns can increase our understanding of the spread of emerging and re-emerging infectious diseases. Only after the normal contact patterns have been accurately modeled, can the simulations predict the impact of behavioral changes on the spread of a pathogen.

These high-fidelity models based on the structure of interactions among individuals can then investigate the effectiveness of different behavior changes, from reducing specific types of contacts to reducing susceptibility and infectiousness though hand washing, wearing protective masks, avoiding crowded places, and school closures. Biased mixing patterns reduce the spread of disease. Without accurate mixing patterns, mathematical models run the risk of overestimating the spread of an epidemic.

5 Acknowledgments

This research has been supported at Los Alamos National Laboratory under the Department of Energy contract DE-AC52- 06NA25396 and a grant from NIH/NIGMS in the Models of Infectious Disease Agent Study (MIDAS) program.

References

1. Ackerman, E., Elveback, L., Fox, J.: Simulation of infectious disease epidemics. Charles C Thomas (1984)
2. Addy, C., Longini Jr, I., Haber, M.: A generalized stochastic model for the analysis of infectious disease final size data. Biometrics 47(3), 961–974 (1991)
3. Anderson, R., Medley, G., May, R., Johnson, A.: A preliminary study of the transmission dynamics of the Human Immunodeficiency Virus (HIV), the causative agent of AIDS. Mathematical Medicine and Biology 3(4), 229 (1986)
4. Barrett, C., Eubank, S., Smith, J.: If smallpox strikes Portland... Scientific American 292(3), 54–61 (2005)
5. Beutels, P., Shkedy, Z., Aerts, M., Van Damme, P.: Social mixing patterns for transmission models of close contact infections: exploring self-evaluation and diary-based data collection through a web-based interface. Epidemiology and Infection 134(06), 1158–1166 (2006)
6. Del Valle, S., Hyman, J., Hethcote, H., Eubank, S.: Mixing patterns between age groups in social networks. Social Networks 29(4), 539–554 (2007)

7. Del Valle, S., Stroud, P., Smith, J., Mniszewski, S., Riese, J., Sydoriak, S., Kubicek, D.: EpiSimS: Epidemic Simulation System. Tech. rep., Los Alamos National Laboratory (2006)
8. Del Valle, S., Tellier, R., Settles, G., Tang, J.: Can we reduce the spread of influenza in schools with face masks? American Journal of Infection Control 38(9), 676–677 (2010)
9. Edmunds, W., Kafatos, G., Wallinga, J., Mossong, J.: Mixing patterns and the spread of close-contact infectious diseases. Emerging Themes in Epidemiology 3(1), 10 (2006)
10. Edmunds, W., O'Callaghan, C., Nokes, D.: Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. Proceedings of the Royal Society of London. Series B: Biological Sciences 264(1384), 949 (1997)
11. Eubank, S., Guclu, H., Kumar, V., Marathe, M., Srinivasan, A., Toroczkai, Z., Wang, N.: Modelling disease outbreaks in realistic urban social networks. Nature 429(6988), 180–184 (2004)
12. Glasser, J., Taneri, D., Feng, Z., Chuang, J., Tull, P., Thompson, W., McCauley, M., Alexander, J.: Evaluation of targeted influenza vaccination strategies via population modeling. PloS ONE 5(9) (2010)
13. Hens, N., Goeyvaerts, N., Aerts, M., Shkedy, Z., Van Damme, P., Beutels, P.: Mining social mixing patterns for infectious disease models based on a two-day population survey in Belgium. BMC Infectious Diseases 9(1), 5 (2009)
14. Hyman, J.M., Li, J.: Disease transmission models with biased partnership selection. Applied Numerical Mathematics 24(2-3), 379–392 (1997)
15. Koopman, J., Simon, C., Jacquez, J., Park, T.: Selective contact within structured mixing with an application to HIV transmission risk from oral and anal sex. Lecture Notes in Biomathematics 83, 316–348 (1989)
16. Longini, I., Halloran, M., Nizam, A., Yang, Y.: Containing pandemic influenza with antiviral agents. American Journal of Epidemiology 159(7), 623 (2004)
17. Mikolajczyk, R., Akmatov, M., Rastin, S., Kretzschmar, M.: Social contacts of school children and the transmission of respiratory-spread pathogens. Epidemiology and Infection 136(06), 813–822 (2008)
18. Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., Edmunds, W.: Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3) (2008)
19. Newman, M., Girvan, M.: Mixing patterns and community structure in networks. Statistical Mechanics of Complex Networks pp. 66–87 (2003)
20. Nishiura, H., Castillo-Chavez, C., Safan, M., Chowell, G.: Transmission potential of the new influenza A (H1N1) virus and its age-specificity in Japan. Euro Surveill 14(22), 19227 (2009)
21. Stroud, P., Del Valle, S., Sydoriak, S., Riese, J., Mniszewski, S.: Spatial dynamics of pandemic influenza in a massive artificial society. Journal of Artificial Societies and Social Simulation 10(4), 9 (2007)
22. Wallinga, J., Teunis, P., Kretzschmar, M.: Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. American Journal of Epidemiology 164(10), 936 (2006)