MicroPox: a Large-scale and Spatially Explicit Microsimulation Model for Smallpox Transmission

Lisa Brouwers

Stockholm Group for Epidemic Modelling
Dept. of Epidemiology
Swedish Institute for Infectious Disease Control
`lisa.brouwers@smi.ki.se`

Dept. of Computer and Systems Sciences
Stockholm University/KTH
`lisa@dsv.su.se`

Keywords: smallpox, epidemiology, transmission, outbreak

Abstract

The motivation for this microsimulation model is the need to analyze and compare effects of implementing potential intervention policies against smallpox. By including contact patterns and spatial information extracted from governmental registers, we strive to make the simulation model more realistic than traditional transmission models, and thus better suited as a policy tool. MicroPox, the probabilistic large-scale microsimulation model described in this article uses real but anonymized data for the entire Swedish population. Since the unique data set contains family relations and workplace data for all Swedes, we have been able to incorporate many of their close social contacts, the type of contacts that are important for the transmission of smallpox. The level of detail of the data makes it possible to capture a large amount of the heterogeneity of the contact structure; most people have a small number of contacts, while a few have a large number. As the data set also contains geographic coordinates for all workplaces and dwellings, we were able to make the model spatially explicit. Besides a description of the model, the article also describes a preliminary experiment in which 50 initially infected persons spread the disease among 2,500,000 people, mainly located in Stockholm, Sweden.

1   INTRODUCTION

Fear of bioterror attacks has forced policymakers worldwide to examine their countries' emergency plans. A common scenario is that an infectious disease agent is deliberately released, threatening an entire region with a lethal epidemic. Smallpox is often cited as an example in such a scenario, because it is perceived to be easily transmitted, it has a high mortality rate, and the level of immunity in the population is low. The task at hand for policymakers is to evaluate potential control and prevention strategies with regard to effectiveness, cost, and risk.

Epidemiological models are often used to estimate the course of an epidemic with and without counter measures, such as, for example, vaccination or isolation of infectious persons. Most epidemiological models are elaborations of the fundamental SIR model, in which the closed and homogenous population is divided into groups (typically Susceptible, Infective, and Removed). A number of differential equations dictate the deterministic flow between these groups over time. In models of this type, all contacts between two individuals from different groups are equally probable, an abstraction that removes the explicit contact structure from the population. Such models have been successfully used as transmission models for highly contagious diseases such as measles, for which the contact pattern of possible contagious contacts very much resembles random mixing. They have shortcomings, however, when it comes to modeling less contagious diseases [Liljeros et al., 2003].

A number of smallpox transmission models have been built to support decision-makers in the task of comparing potential control strategies. An article by Edward Kaplan and colleagues has received much attention [Kaplan et al., 2002]. The article describes their model, based on a set of coupled differential equations, and a number of experiments in which different control strategies are applied to a scenario in which smallpox is spread in the population by terrorists. From the results of the experiments, the authors recommend mass vaccination as the best policy intervention in case of a terrorist attack. Note that vaccination is always associated with risk: according to historical data, post-vaccinial encephalitis has occurred at a rate of 1 per 300,000 vaccinations, of which 25 percent were fatal [Henderson et al., 1999, Henderson 1999]. A fraction of the population would experience an adverse event if a mass vaccination program were carried out, and as many as 10 million individuals in the U.S. may be immunocompromised and at risk for severe complications. To minimize the danger, these high-risk individuals and their close contacts would have to be excluded from the vaccination program, at the cost of lowering the level of immunity in the population [Kemper and Matthew 2003]. The same year as Kaplan's model was presented, Samuel Bozzette [Bozzette et al., 2003] presented a stochastic transmission model from which they demonstrated that mass vaccination is recommendable only if the probability of a severe bioterror attack is high. Both models assume random interaction, implying that transmission between two individuals living 1000 kilometers apart is just as likely as transmission between

family members. In contrast with this assumption, several scientists have pointed out that the attack rates between close contacts and casual contacts differ; the probability for transmission is much higher if the contact is prolonged and close [Fenner et al. 1988; Fenner 1988]. Elizabeth Halloran has proposed another policy: post-outbreak ring vaccination. Their simulation model [Halloran et al., 2002] involves a 2000-person community that is structured into families with 1 to 7 members. Transmission occurs in the household and at places where people meet, such as in neighborhoods and at schools.

The model only includes eight instances of places of this type; all high-school children attend the same school and the entire community is divided into four neighborhoods with 500 people each.

Data from historical outbreaks supports the view that smallpox is primarily spread between close contacts, as reported by Eichner and Dietz, who have analyzed data from an outbreak in Nigeria [Eichner and Dietz 2003]. In 79.9 percent of the cases, the disease was transmitted between members of the same household, and in as many as 93.3 percent of the cases, the transmission occurred between persons belonging to either the same household or to the group of "other close contacts." These findings point to the importance of including reasonable estimations of the contact patterns.

An increasing number of models include a more ambitious and more realistic representation of the population. One early example of a microsimulation model (a model that represents each individual explicitly) is the GERMS model [Adams et al., 1999]. More recent examples are a model based on parameter estimations from the Nigeria outbreak [Eichner 2003] and a model in which the transmission probabilities vary with the type of contact, close or casual [Kretzschmar et al., 2004].

In an individual-based and spatially explicit model of Portland, in the United States, a social network was constructed over a synthetic population [Eubank et al., 2004]. The network is based on land-use data and census data on transportation in the city. The simulation approach is bottom-up; no value for the basic reproduction rate R_0[1], is given prior to the simulation, but instead is derived afterwards.

2   THE MICROPOX MODEL

The data set we use for this transmission model is unique. It belongs to the Spatial Modelling Centre (SMC) and is delivered to them by Statistics Sweden (SCB). SMC has been granted permission to use the data set for model research. Since we collaborate with Einar Holm and Kalle Mäkilä at SMC in this project, we may use a subset of this data set for this specific transmission model. This

cooperation gives us access to real but anonymized data from 1998 for the entire Swedish population, consisting of close to nine million individuals. The database contains references to the workplaces of each person. All dwellings and workplaces have geographic coordinates, and each coordinate maps to one cell (100 x 100 meters) in a grid that represents Sweden.

The simulation model is dynamic; the simulation proceeds hourly. Each day is divided into three 8-hour periods. In the simulation, morning people make probabilistic choices about where to go. First they check their health; if they are well and have a job they go to their workplace. If they are ill, unemployed, or retired they stay at home. After eight hours at work all persons return home and remain there until the next morning.

Persons and places are the basic entities of the MicroPox model. Since transmission only occurs where they are co-located, the transmission process executes at particular places, and we have distinguished the most important types of places for smallpox transmission. These include dwellings, kindergartens, schools, hospitals, and offices. Furthermore the model includes two abstract places, "travel" and "neighborhood." More information about the different places is found in the section "Representation of Places."

A simulation is run for a number of days, normally 100 to 300. Three groups of parameters are set before the start of a simulation:

- Simulation parameters
  - Number of days
  - Number of replications
  - Names of files to be read (population, etc.)
- Model parameters
  - Number of initially infected persons
  - Use hubs (here, a hub is an individual with a very large number of contacts)
  - Transmission probability (per *place*)
  - Stage Weight (prodromal, symptomatic I, symptomatic II)
  - Probability that a person in the prodromal stage feels well
- Underlying assumptions
  - Maximum number of contacts (per place)
  - Daily probability that a person is ill (from other cause than smallpox)
  - Daily probability that a person who is ill visits an emergency ward
  - Daily probability that a person travels

The attributes are described in the following five subsections and their use is exemplified in the last section, "Experiment." During the simulation, all transmission events are logged and written to a file, called the event file. It contains the time (day from start of simulation) of the transmission, at what type of place the transmission occurred, person-id of the infected and the infector, and the age and the sex of the infected. This explicit logging makes

---

[1] R_0 defines the average number of persons who become infected by one infector in a totally susceptible population if no policy interventions are implemented.

it possible to follow the chain of transmission all the way back to the initial infector. It also makes it possible to plot the locations where at least one infected individual lives (see Figure 1). Since the coordinates are plotted daily, the spatial spread of the epidemic is captured.
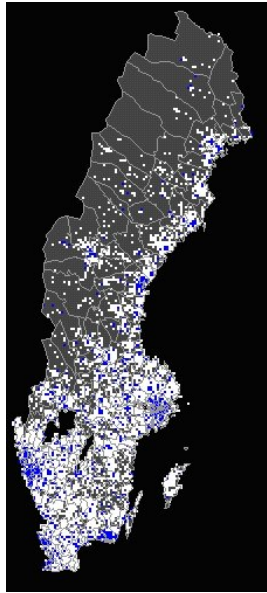

Figure 1. Visualization of spatial distribution of all infected individuals (snapshot). The visualization tool was developed by Kalle Mäkilä and Julien Gaffuri at SMC.

2.1 Representation of Individuals

Each individual has the following attributes: *person-id, sex, age, family-id, workplace, immune, state,* and *clinic*. Persons who are unemployed or retired do not have a workplace. Children spend their days at school or at kindergarten, which may be seen as their "workplace." The data set contains schools and kindergartens in the form of workplaces for adults, but it contains no information about which school a child is enrolled in. Therefore, we must generate a proxy for this connection in the MicroPox model. The choice of kindergarten/school/college/university depends on age and on the geographic distance between school and dwelling. If the closest appropriate type of school/kindergarten is not yet filled, the child is added to its enrollment list and the workplace attribute is updated accordingly. All children from 1 to 15 are connected with a school or kindergarten. For young people 16 to 20, we first check whether they have a workplace, and if not the majority of the unemployed youths are placed in schools.

The attribute *immune* tells whether an individual is immune to smallpox or not. Immunity comes either from vaccination or from having survived a smallpox infection. The *state* attribute describes what stage of smallpox the person currently is in. An uninfected but susceptible person has state 0 for instance (see subsection "Representation of the Disease" for more information about stages and transition rules). *Clinic* holds information about which emergency ward the individual should visit if needed. The choice of clinic is based on vicinity.

2.2 Representation of Places

All places have the following attributes: *type, transmission probability, current infection risk, coordinates,* and *members* (a list). *Dwellings* are places where families live. They have the additional attribute *family-id*, which makes it is possible to connect families and dwellings. Several dwellings can have the same coordinates, since the cells are 100 square meters in size. If the *member* list of the dwelling contains one or more infectious people, a neighborhood list is created. *Emergency wards* are specializations of the *hospital* place. Some hospitals have an emergency ward. Emergencies have a *patient list*, representing patients in the waiting room. At the time of transmission, the *patient list* is concatenated with the *member list* (staff) to enable transmission from one group to the other. *Office* is the default place type for workplaces.

The last two places can be seen as proxies for relations that are important to include in the social network, but for which we have no real data. The purpose of the *travel* place is to represent the contacts between persons away from their ordinary location. The proxy represents one-day trips within Sweden. Each region in Sweden is provided with a *travel* place. At this place people who travel to this region (or within the region) will meet other travelers. A fraction of the population travels each morning; the destination of their travels is probabilistically determined by using a gravitation model. The number of people in the region and the vicinity from the dwelling of the traveler decide the probability that the trip will have that region as its destination. Short trips are more likely than long ones, and trips to a densely populated region are more likely than to one that is sparsely populated. *Neighborhood* is a proxy for random encounters. These encounters could take place at grocery stores, cinemas, public transports or other places where many people meet. The underlying assumption for this is that it is more likely for a person to meet someone from his or her immediate area than from far away. When a transmission has occurred in a dwelling, a *neighborhood* list is created and filled with a number of individuals. The probability that a person will be added to this list decreases with distance.

2. 3 Representation of the Disease

For an overview of the clinical features of smallpox, consult an overview, for instance by Fenner [Fenner 1988]. However, there is no general agreement upon the exact values of the disease parameters such as time and duration of infectiousness. We have divided the disease period into six stages:

- Incubating I: asymptomatic, vaccine sensitive, not infectious
- Incubating II: asymptomatic, vaccine insensitive, not infectious
- Prodromal: acutely ill, possibly infectious

- Symptomatic I: overtly symptomatic, highly infectious
- Symptomatic II: overtly symptomatic, infectious
- Dead or immune: all recovered persons become immune

The transition from one stage to the next is time-dependent: after a number of days in one stage the infected person enters the next stage. The duration is decided either deterministically or probabilistically.

2.4 Smallpox Transmission

The first step in the transmission process is to count the number of infectious members at each place. The *current infection risk* of each specific place depends on:

1. the predefined *transmission probability* for that type of place, *TrP*
2. a vector of *stage weights,* **SW,** for the three infectious stages
3. a vector containing the number of infectious people at each of the three specific stages of infectiousness **NoInf**
4. the number of members at that specific place, *NoC*

The risk that a susceptible individual will become infected by infectious individuals at a certain place can be written as

$$P(infected|place)_t = 1 - \prod_{stage=1}^{3}((1 - TrP_{place} \cdot SW_{stage})^{NoInf_{stage}}) \quad (1)$$

Equation 1. We calculate the probability of being infected at a certain place at a certain time by deducting it from the complement of the probability of *not* being infected in all three stages (1=prodromal, symptomatic I, and symptomatic II.

For places with a large number of members it seems unrealistic that all persons meet during a day. If the number of people at a place exceeds the value of the parameter *maximum number of contacts, NoC,* for that type of place, the *current infection risk* from each single infectious individual is reduced by the proportion δ of the total number of members, *NoM* at the place, as described in equation 2.

$$P(infected|place)_t = 1 - \prod_{stage=1}^{3}((1 - TrP_{place} \cdot SW_{stage} \cdot \delta)^{NoInf_{stage}}),$$
$$\text{where}$$
$$\delta = \begin{cases} 1 & :NoM \leq NoC \\ \frac{NoC}{NoM} & :NoM > NoC \end{cases} \quad (2)$$

Equation 2. The probability that a susceptible member will be infected by infectious member(s) where the number of members at each place may exceed the maximum numbers of contacts each day.

2.5 Initially Infected

The number of *initially infected* persons can vary from 1 to the size of the entire population. In a bioterror scenario, a credible number would be somewhere between 1 and 100.

There are two ways of choosing which individuals to infect: randomly or by identifying 'hubs'. A hub is a person who has extremely many contacts in the social network. He or she may have a large family and work at a large workplace. The program makes it possible to decide limits for minimum number of family members and for members of the workplace. If the value of the parameter *use hubs* is 1, a list of hub-persons is created from which the initially infected persons are picked.

3 EXPERIMENT

We have performed a small experiment to demonstrate how the model can be used. We used a population that consists of all persons in Stockholm plus 1/10 of the persons from other places in Sweden (2,500,000 persons in total). The simulation was executed on a PC with 2 Gb RAM memory. The following parameter values were used:

Table 1. Parameter values that were used in the simulation.

Parameter	Value
Number of days	250
Number of replications	1
Number of initially infected	50
Use hubs	0 (no)
Transmission probabilities (places):	
Dwelling	0.25
Kindergarten, School, Hospital	0.15
Office	0.05
Travel	0.15
Emergency ward	0.25
Neighborhood	0.10
Maximum number of contacts:	
Dwelling	20
Kindergarten, School	30
Hospital	50
Office	20
Travel	30
Emergency ward	50
Neighborhood	30
Daily probabilities of:	
Feeling well in prodromal	0.05
Feeling ill (not smallpox)	0.1
Visiting the emergency if ill	0.1
Traveling	0.1
Stage Weights (for transmission)	
Prodromal	0.2
Symptomatic I	1.0
Symptomatic II	0.5

We varied the level of infectiousness by the stage. Reducing infectiousness during the prodromal stage drastically reduced the potential for a person to infect others since infected persons stay home as soon they become symptomatic. We assumed the level of residual immunity in

the population to be zero; no persons were immune at the start of the simulation.

Table 2. Duration of the smallpox stages.

Stage	Duration	Distribution
Incubating I	3	Deterministic
Incubating II	4-16	See Fig. 2
Prodromal	3-5	Uniform
Symptomatic I	4	Deterministic
Symptomatic II	16	Deterministic
Immune		2/3 survive
Dead		1/3 die; prodromal stage + 7--14 days; uniform distribution

The first stage, when a person is vaccine sensitive was set to 3 days. Total length of the period from infection to onset of fever varied between 7 and 19 days. Fig. 2 presents the probability density for the duration of this period (Incubating I plus Incubating II).
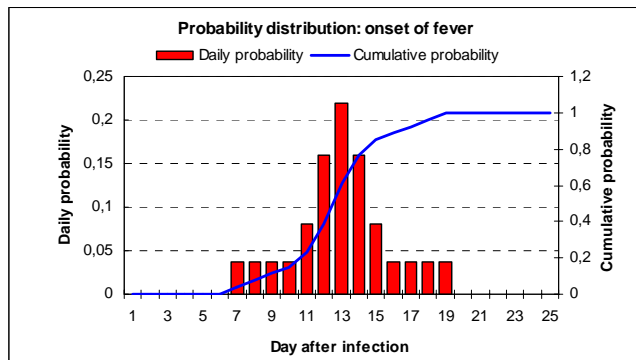
Figure 2. Daily probabilities for transition to state Prodromal. Time of infection is day 0.

The value 0.05 of the parameter *Feeling well in prodromal* guarantees that most persons fall ill and stay home when they enter the prodromal stage. The remaining 5 percent can infect persons away from home. On the last day of the prodromal stage all persons visit the emergency ward located closest to their dwelling, where they may infect other patients or hospital staff. When the rash appears on a person (Symptomatic I) he or she stays home and becomes highly infectious. This period of early rash lasts for 4 days. Symptomatic II stage consists of three periods: pustular rash, pustules and scabs, and resolving scabs. A person is infectious throughout this stage. The chance of survival is 2/3. The figure corresponds to the mortality rate used in other smallpox models; however, it is possible that it is to pessimistic if recent years' advances in health care are considered. If a person doesn't survive, he or she dies one to two weeks after entering the prodromal stage, the distribution within the time interval is uniform. Persons who survive the disease become immune after leaving stage

Symptomatic II. No policy interventions were used in the experiment, that is, the infected person was neither isolated nor vaccinated, and no contact tracing was performed. As shown in Figure 3, it took quite long time before the epidemic took off.
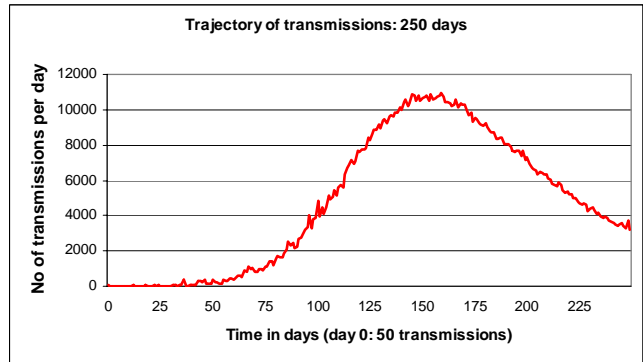

Figure 3. Daily number of newly infected cases over the entire simulation period (250 days).

Including the 50 initially infected persons 1,190,121 infections occurred (Figure 4).
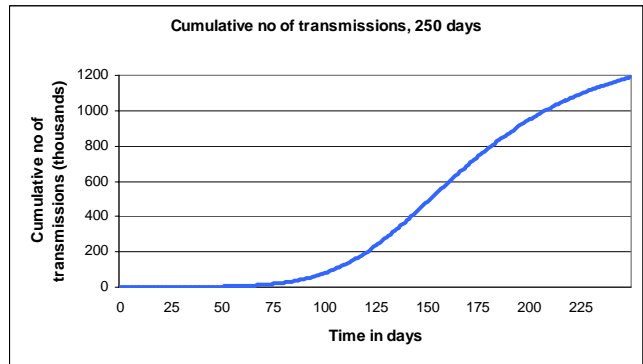

Figure 4. Cumulative number of infections.

By analyzing the *event file* it is possible to study transmission in greater detail. Figure 5 shows the number of persons in each stage per day.
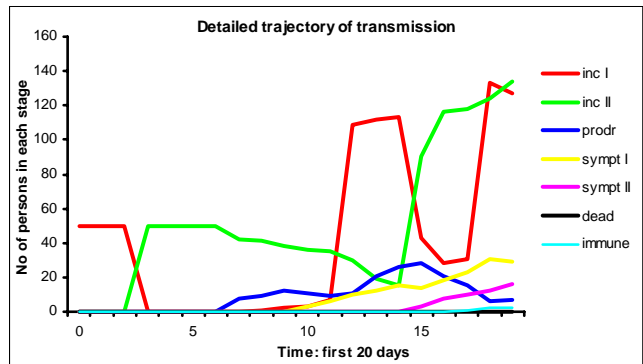

Figure 5. Number of persons in different stages, first 20 days.

Since we log where transmissions occur, it is possible to analyze the distribution of places for transmission events.

Figure 6 shows that in this experiment most transmissions took place in dwellings.
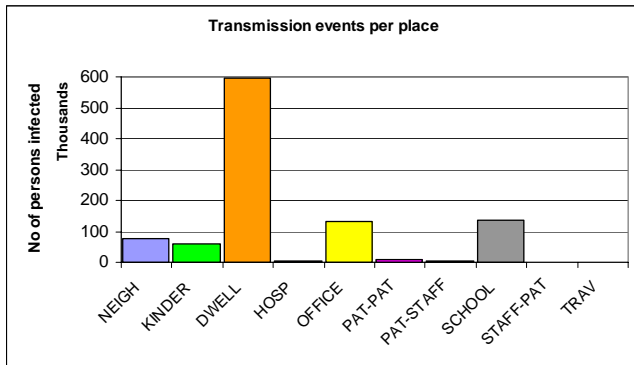

Figure 6. Total number of transmission events per place: *Neighborhood, kindergarten, dwelling, hospital, office: patient infects patient (emergency), patient infects staff (emergency), school, travel; staff infects patients (emergency), travel*

4 DISCUSSION

The experiment presented is merely included to illustrate the features of the model; since the model is still under development it has not yet been fully validated. The next step for this model is to include policy interventions, starting with isolation and contact tracing. We will also introduce a level of *national smallpox awareness*. Before the first case is verified, the level will be zero. After one or more verified cases, the routines at emergency wards will be changed; infected patients will be redirected to special infection clinics, etc. At some *awareness* level the daily habits of persons will change as well; people keep their children home from school and reduce their social contacts and travel less.

REFERENCES

Henderson, D.; T. Inglesby; J. Bartlett; M. Ascher; E. Eitzen; P. Jahrling; J. Hauer; M. Layton; J. McDade; M. Osterholm; T. O'Toole; G. Parker; T. Perl; P. Russell; K. Tonat. 1999. "Smallpox as a biological weapon - Medical and public health management." *JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION 281*, No: 22, JUN 9: 2127-2137.

Henderson, D. 1999. "Smallpox: clinical and epidemiologic features." *Emerging Infectious Diseases 5,* (537-539).

Liljeros, F; C. Edling; L. Amaral. 2003."Sexual Networks: implications for the transmission of sexually transmitted infections." *Microbes and Infection 5,* (189-196).

Kaplan, E. H. and D. L. Craft; L. M. Wein. 2002. "Emergency response to a smallpox attack: The case for mass vaccination." *Proceedings of the National Academy of Sciences of the United States of America 99*, No: 16, Aug 6: 10935-10940.

Kemper, A. R. and M. D. Matthew. 2003. "Under scrutiny: Smallpox vaccine recommendations." *Expert Opinion on Pharmacotherapy 4*, No: 8, (1207-1214).

Bozzette, S.; R. Boer; V. Bhatnagar; J. Brower; E. Keeler; S. Morton; M. Stoto. 2003. "A model for a smallpox-vaccination policy." *NEW ENGLAND JOURNAL OF MEDICINE 348*, No: 5, JAN 30: 416-425.

Fenner, F.; D. Henderson; I. Arita; Z. Jezek; I. Ladnyi. 1988. "Smallpox and its eradication." World Health Organization, Geneva.

Fenner, F. 1988. "The eradication of smallpox." *IMPACT OF SCIENCE ON SOCIETY 38*, No: 2, 147-158

Halloran, M.; I. Longini; A. Nizam; Y. Yang. 2002. "Containing bioterrorist smallpox." *SCIENCE 298*, No: 5597, NOV 15: 1428-1432.

Eichner, M. and K. Dietz. 2003. "Transmission potential of smallpox: Estimates based on detailed data from an outbreak." *AMERICAN JOURNAL OF EPIDEMIOLOGY 158*, No: 2, JUL 15: 110-117.

Adams, A. L.; J. S. Koopman; S. E. Chick; P. J. Yu. 1999."GERMS: an epidemiologic simulation tool for studying geographic and social effects on infection transmission." In *Proceedings of the 31st Conference on Winter Simulation: Simulation---a Bridge to the Future* (Phoenix, Arizona, United States. ACM Press, City, (1549-1556).

Eichner, M. 2003. "Case isolation and contact tracing can prevent the spread of smallpox." *AMERICAN JOURNAL OF EPIDEMIOLOGY 158*, No: 2, JUL 15: 118-128.

Kretzschmar, M.; S. van den Hof; J. Wallinga; J. van Wijngaarden. 2004. "Ring vaccination and smallpox control." *EMERGING INFECTIOUS DISEASES 10*, No: 5, MAY: 832-841.

Eubank, S.; H. Guclu; V. Kumar; M. Marathe; A. Srinivasan; Z. Toroczkai; N. Wang. 2004. "Modelling disease outbreaks in realistic urban social networks." *NATURE 429*, No: 6988, MAY 13: 180-184.

Holm, E.; K. Holme; K. Mäkilä; M. Mattson-Kauppi; G. Mörtvik. 2002. "The SVERIGE spatial microsimulation model: Content, validation, and example applications." In *GERUM kulturgeografi*. Umeå, Sweden: CERUM Kulturgeografi, Kulturgeografiska Institutionen/SMC, Umeå University.