I. Troch, F. Breitenecker, eds. ISBN 978-3-901608-35-3

LONG TERM BEHAVIOR OF AGENT BASED EPIDEMIC SIMULATION OF STREPTOCOCCUS PNEUMONIAE - A MATHEMATICAL STATUS REPORT

Florian Miksch[1], Nikolas Popper[2], Günther Zauner[2], Irmgard Schiller-Frühwirth[3], Gottfried Endel[3]

[1]Vienna University of Technology, Institute for Analysis and Scientific Computung, Vienna, Austria;
[2]"Die Drahtwarenhandlung"-Simulation Services, Vienna, Austria;
[3]Evidence Based Economic Healthcare, Hauptverband der Österreichischen Sozialversicherungsträger

Corresponding Author: Florian Miksch, Vienna University of Technology

Abstract. Streptococcus Pneumoniae is a bacterium that causes several infections like pneumonia, otitis media, meningitis and sepsis. Today, due to vaccination against many other epidemics, these are one of the most common and dangerous sicknesses for small children in developed countries. The purpose of this work is to develop an agent based model that is able to simulate a pathogen like Streptococcus Pneumoniae and to understand the long term effects of different vaccination strategies. The main part of the work is to explain the model structure in detail, understand the benefits and problems of such an agent based epidemic model and find out how the model behaves in different situations. Also profound data search has to be done to understand the whole system and identify the model parameters as accurately as possible. It is important to realize how to cope with restrictions in data quality and different results of different studies.
Agent based models are only one approach to simulation of epidemics so other people of the research group are simulating the same problem with different model types. In the end the results of the different models should be compared and benefits of the different approaches have to be discussed.

1 Streptococcus Pneumoniae and the vaccination

The standard therapy against Streptococcus Pneumoniae is Penicillin but due to wide spread of the pathogen and its combat Penicillin resistance grew especially in North America [2].

In the early eighties a vaccination against Streptococcus Pneumoniae was invented. The problem with this vaccination is that it provides no protection for children younger than 24 months because it does not produce an immune reaction at them. But there are two reasons why it is quite important to protect young children: First, young children (as well as old people) have the by far highest disease rate because their immune system is quite weak compared to adults. And second, young children are suspected to be responsible for most of the transmissions of Streptococcus Pneumoniae.

The real world system seems to be very complex; additionally many different medical studies report different results. Fact is that a significant part of the population is infected with Streptococcus Pneumoniae and can infect other people while only a small part falls sick with it. All the other people get rid of it after a while and do not even realize that they have been infected. All together there are more than 90 serotypes of Streptococcus Pneumoniae. It seems to be clear that the prevalent serotypes and the percentage of infected individuals in a population is regional different and changes over time.

The PCV-7 vaccination provides protection from 7 frequent serotypes, in North America it covers about 90% and in Europe about 70% of the prevalent serotypes [2].

1.1 Hospital data

Comprehensive hospital data from Austrian hospitals about pneumonia, meningitis and sepsis is available. This data can be normalized using population data, and then some statements can be made about the probability that a random person with a certain age falls sick with one of these sicknesses. Therefore the number of infected people is a black box because there are no corresponding studies available for Austria.

The important result from this data research is that the probability to fall sick is very high for babies, then it gets lower for children, it stays on a very low level for adults and starts to increase dramatically again from age 65 onwards. It is obviously that pneumonia is by far the most common sickness caused by Streptococcus Pneumonia. Meningitis and sepsis are only occasional instances all over Austria.

1.2 Additionally there are two considerable phenomena that should be examined in detail:

Serotype replacement. One problem is that there are many uncommon serotypes that are not covered by the vaccine. Some studies presume that vaccinated people get infected with uncommon serotypes instead so other serotypes could become more common. So the real long term effect of the vaccination could be much lower than expected.

There is even one clinical study that reports serotype replacement [3]. But usually serotype replacement cannot be observed because at the moment only a too small part of the population is vaccinated and therefore random effects cannot be separated. The few vaccinated people really benefit in many studies because they just do not get in touch with uncommon serotypes and compared to the rest of the population they are not many enough to make uncommon serotypes more common.

This model should be able to give some more information if such a phenomenon can be expected and how strong it might be.

Herd immunity. Another presumption is that there exist substantial effects for the whole population if only children are vaccinated.

The explanation is on the one hand a much higher percentage of children than adults are infected, on the other hand in some studies it is presumed that most transmissions of this pathogens are between infants and adults because their physical contact is often much closer than between two adults.

Missing Parameters. First, there are many studies with controversial results about the percentage of the children infected with Streptococcus Pneumoniae. But no data about the part of infected adults, especially old people, can be found. In Austria only the reported cases of disease outbreaks with admission are reported. Second, clinical studies do not provide data about what part of infected people fall sick within a defined period or, at least, how many infected people are sick at a certain time.

2 The Model

The model is an agent based model with some special functionality. Requirements are:

- Simulate over a long period (approximately two decades) to find out long term effects
- Consider a changing population structure because of long term simulation
- Implement a social model to simulate contacts between individuals
- Simulate more than one pathogen to differentiate between covered and non-covered serotypes by the vaccination

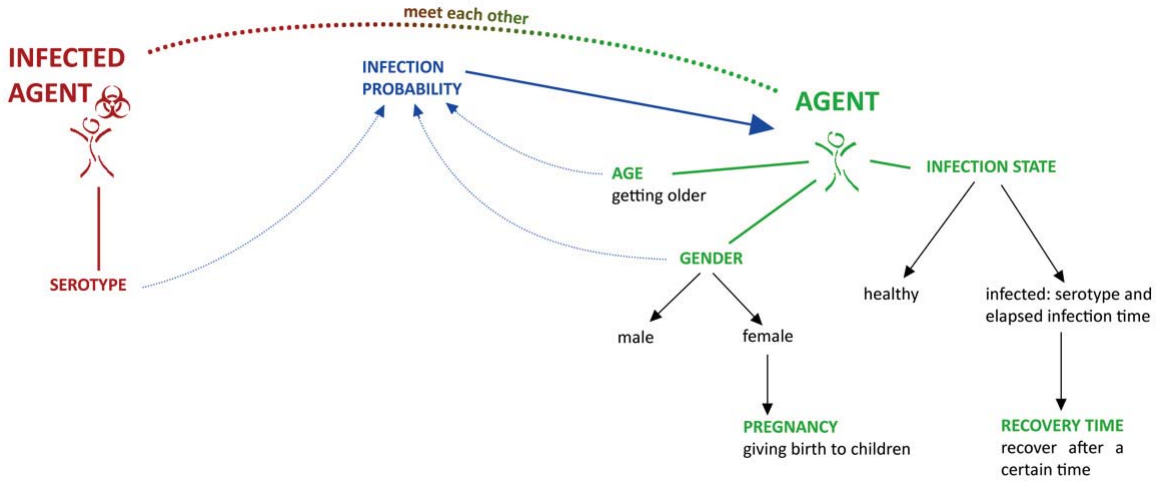

Figure 1: A graphical description of the agent based model with main focus on the real social system behavior

The model consists of three parts: The population part, the social part and the epidemics part.

The population part. Single persons have the following attributes: Age, Gender, Infection State and Pregnancy (women only). Additionally the auxiliary attributes unique ID-Number for identifying a single person, Age Class and Infection Protocol is stored. Changing population is realized so that people are getting older, they can die and women can give birth to babies. The parameters can be easily identified with real population data of Austria provided by "Statistik Austria" [4].

The social part. Epidemics can be spread only through direct contacts between two persons. Contacts can happen in a household, at work, while meeting friends and randomly. It is not possible to simulate such a detailed system efficiently so it is modeled in a simplified way based on a suggestion in an US-American paper [2]:

Consider a connection between two persons only without mentioning how and where they meet. A connection between two persons means that these persons meet each other in that time step. Two people are called friends as long as a connection between them exists. The social net is never constant so it has to change from time to time. At first some constants have to be defined: The average number of connections per person, the break-up rate and

the connection rates. Every time step some of the connections depending on the break-up rate are deleted and new connections, depending on connection rates, are added until the average number of connections per person is reached. The new connections are, depending on the connection rates rate, partly completely random, partly two people of the same age group and partly two random "friends" of a person who are not connected yet.

The epidemics part. For the epidemics part two constants have to be defined: The infection probability and the recovery time.

The procedure of the epidemics part is three steps: At first process all connections between an infected and a healthy person one by one and let the healthy person become infected with the infection probability which is dependent on the age and gender of the susceptible person and the serotype. In the second step increase the infection time by one of all infected persons. This value is stored in the Infection State attribute of a person and means how long an infected person has been infected already. At the end let all infected persons recover that have reached recovery time which means to set their Infection State the healthy state.

Additional functionality:

- Simulating two or more pathogens where a single person cannot be infected with more than one pathogen at one time.
- Individual infection probabilities and recovery time depending on the age and gender of the person and on the pathogen.

3 Behavior of the model

3.1 Simulating one serotype

While simulating one pathogen only the model behaves very predictable. Depending on infection probability and recovery time the number of infected and healthy persons reaches, independent of the start values, a constant level. A higher infection rate, a longer recovery time and a lower number of connections per person cause a higher level of infected people. Additionally a correlation between two parameters can be observed: Doubling (or halving) the infection rate has the same effect as halving (or doubling) the number of connections per person. Correlations between the recovery rate and other parameters cannot be found but it has another effect: A higher recovery rate causes a more inert system which means that overshooting occurs for long recovery times but not for shorter ones.

The results seem to be really good and useable because the model is predictable and the parameters can be adjusted easily and efficiently. The prevalence of overshooting shows that the model does not correspond with the first-order differential equations of the simplest SIS-model. So differential equations of order two or higher have to be considered if the model should be identified with a differential equation model.

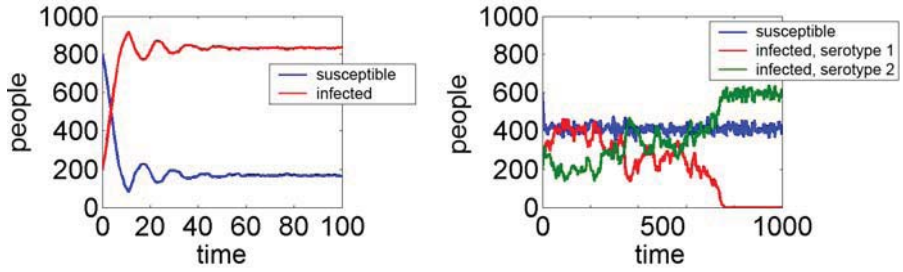

Figure 2. Representative results of the model. On the left side a result of a single serotype run where overshooting can be observed and on the right side a result of a run with two simulated serotypes.

3.2 Simulating two serotypes

Things become more difficult when simulating two serotypes. If both serotypes are equally strong (that means that they have the same parameters) the behavior will be completely randomly until one of the serotypes became extinct, then the system continues in the stable one-serotype system.

If one serotype is stronger (higher infection probability or recovery time) the weaker serotype will become extinct very fast. Such behavior does not correspond with reality at all. There are over 90 different serotypes and all of them survive even if they only occupy a very small part of all infected people. At least all the time the numbers of healthy people and infected people altogether stays on a constant level. The level is the same as the level of the one-serotype model. If the two serotypes have different parameters the number of healthy people in the two-serotype model is the same as in the one-serotype model with the parameters of the stronger serotype (the stronger serotype is the one that will survive).

3.3 The updated model

Because of the strange behavior another approach for the infection part is developed. In the original model only healthy persons are susceptible. This means that an infected person can only be infected with another serotype after he or she has recovered from the current infection.

Clinical studies report that sometimes infected persons get infected with another serotype without being recovered meanwhile. So the next try is to allow infected persons to get infected with another serotype and loose the old serotype simultaneously before the end of the recovery time.

Including this rule into the model structure it is now possible to simulate two different serotypes within a more stable system. First simulations show that it is not as predictable as the one-serotype system but if the parameters are chosen properly serotypes do not become extinct any more. Also a tendency of how strong the single serotypes are can be observed very well. This makes it possible to give proper statements about the behavior of two or more serotypes in one system now and additional results about serotype replacement can be acquired.

4 Results

Results are very controversial. On the one hand we have predictable results – the number of infected people together. On the other we have a partly chaotic and non-predictable system. Things got better with the updated model which is not completely tested yet. Anyway the results of our simulation runs show the same problem than in real life. Different studies in different countries report completely different results, so it seems that the real system is not stable in a local area at all.

Simulation of vaccination strategies. Extensive testing with different parameters of vaccination strategies has not been done yet. First results show a good protection from vaccinated serotypes even if only a part of the population is vaccinated. But a very strong serotype replacement can be observed too.

Simulation runs based on real data. The problem is that many different clinical studies from Europe and the USA show completely different results. At first it has to be specified which real datasets we consider as our reality for the model, then it is possible to determine the model parameters and simulate it.

Simulation of herd immunity. No data about infection rates for different ages, especially for adults, could be found yet for Austria, so the model has not been tested on herd immunity so far. Once we simulate the model with parameters based on real data we will get results for herd immunity automatically.

5 Outlook

Finish testing of the updated model. When finished, simulate the model with different parameters corresponding with clinical studies and find out effects of vaccination of small children.

Up to now all runs are done with same parameters for all age groups regarding infection rates. As there is no data about infection rates of adults and old people available, these parameters could be estimated using sickness probabilities for different ages that are calculated from hospital data.

With consensus for Austrian data about infection, transmission and sickness rates, vaccination effects on individuals for different age groups can be simulated and we will receive stable and reliable results for serotype replacement ad herd immunity with the updated infection model.

6 References

[1] Emily M. Jin, Michelle Girvan, M. E. J. Newman (2001):
Structure of growing social networks, Physical Review E, Volume 64, 046132

[2] Christa Claes, Ralf René Reinert, Christoph Vauth, Wolfgang Greiner (2008):
Heptavalenter Pneumokokkenkonjugat-Impfstoff (PCV7)

[3] D. Bogaert, R. H. Veenhoven, M. Sluijter, W. J. W. Wannet, G. T. Rijkers, T. J. Mitchell, S. C. Clarke, W. H. F. Goessens, A. G. Schilder, E. A. M. Sanders, R. de Groot, P. W. M. Hermans (2005):
Molecular Epidemiology of Pneumococcal Colonization in Response to Pneumococcal Conjugate Vaccination in Children with Recurrent Acute Otitis Media
(Journal of clinical microbiology)

[4] Statistik Austria, Population Data (2008):
http://www.statistik.at/web_de/statistiken/bevoelkerung/index.html