EVALUATION OF DIFFERENT MODELING TECHNIQUES FOR SIMULATION OF EPIDEMICS

Günther Zauner[1], Niki Popper[1], Felix Breitenecker[2]

[1]dwh Simulation Services, Neustiftgasse 57-59, 1070 Wien, Austria
[2]Vienna University of Technology, Vienna, Austria

guenther.zauner@drahtwarenhandlung.at (Günther Zauner)

Abstract

The spread of infectious disease and the induced harm and deaths are of main interest in populations since the existence of modern social structures. For example the Spain flu in the early 20th century or the modern swine flu is influencing the overall world social systems.
To find the best suited strategy against an illness taking into account additional boundaries it is necessary to calculate different scenarios in advance. Thereby assumptions regarding bounded resources of vaccines or money for a strategy as well as epidemiological key parameters have to be taken into account.
In mathematical theory the ordinary differential equations of Kermack and McKandrick in 1927 describing the spread of disease are one key step in simulation of the behavior of an epidemic. The so called SIR (Suspected – Infected – Recovered) model.
Nevertheless these equations are not the only way of modeling complex disease behavior. In this work the main solution strategies used in modern pharmacoeconomic output research are discussed. These strategies are Markovian models, ODE models and agent based modeling techniques.
The three strategies are explained briefly and in the second task the differences and benefits as well as problems of each method are figured out. In the outlook a concept for feedback in modeling and more general in problem solving in pharmacoeconomic decision making regarding infectious diseases is presented.

Keywords: modeling techniques, epidemics, evaluation of strategies

Presenting Author's biography

Günther Zauner. He studied Technical Mathematics with a focus on mathematical methods in computer science at Vienna University of Technology. His current field of work is modeling and simulation in health technology assessment (HTA), development of simulation strategies for infectious diseases for dwh Simulation Services. He is also working on a PhD thesis supervised by Prof. Felix Breitenecker in Vienna.

1 Basic Assumptions

A SIR – model is a simple model description for simulation of spread of an illness. Therefore three types of persons are taken into account:

- S – Susceptible persons: These persons are healthy and can become infected by infected persons.
- I – Infected persons: These persons are infected and can during this phase infect healthy persons (type S). The infected persons get healthy again with a predefined probability or after a defined time period. After getting healthy they are immune against the modeled disease.
- R – Recovered persons: This group includes all persons that have been infected and are already convalesced. In this model structure people can not suffer from the same illness more often.

At the beginning of a SIR – simulation normally the main part of a population is assumed to be susceptible. Only few infected persons are given in the beginning; the dynamic change of the fractions is of interest.

SIR – specific parameters are the infection probability, which defines the likelihood that a healthy person gets infected during the contact with an infected person, and the recovery probability, which defines the probability that an infected person (I) gets recovered. This is pointed out by a change of its infection state from I to R.

2 Markovian models

A Markov process is a random process X_t whose future probabilities are determined by its most recent values. For discrete time steps the calculation of X_{n+1} is independent of the knowledge of $X_1, X_2, \ldots, X_{n-1}$, only X_n has to be known and therefore:

$$P[X_{n+1} \mid X_1, X_2, \mathrm{K}, X_n] = P[X_{n+1} \mid X_n]$$

When focusing on discrete time steps the so defined Markovian process is also called Markovian chain.

Markovian models are part of the top – down modeling techniques and are in general representing the behavior of one cohort over the time span of interest. As cohort in a Markovian model in general a group of persons with identical behavior is assumed. The model is not taking into account other persons outside the cohort and their influence into the overall behavior of a population or different parameterization in this group of persons. The probabilities of getting infected, recovered, etc. are based on statistics and often decision trees (general structure shown in Fig. 1) are used to get the values for each time step.

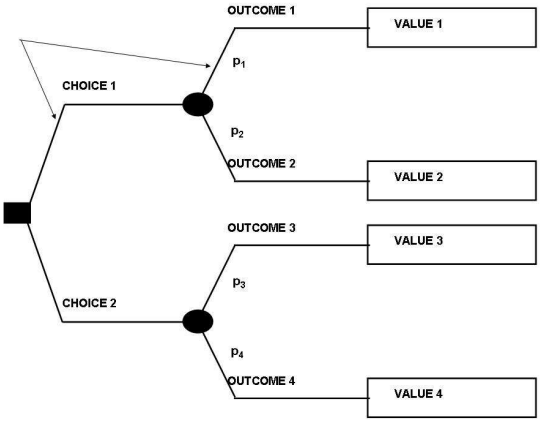

Fig. 1 General structure of a decision tree as used for calculation of the probabilities for one step in a Markovian model. Thereby Choice 1 and Choice 2 can for instance represent the decision between vaccination and no prophylaxis.

As this method is data based and is not modeling the real dynamic of the contact behavior of the population some problems occur:

- The structure of the Markovian model does not allow a deeper insight into the dynamic effects of spreading of the illnesses influenced by vaccination and dynamic feedback due to the properties of the Markovian structure.
- With this kind of model it is not possible to represent dynamical nonlinear effects like herd immunity. This can only be done by fixed parameters which are in general defined by expert knowledge without any possibility of reliability testing or validation.
- Assumptions for heterogeneous population structures cannot be modeled. Especially the influence on other age groups or serotypes is very important to be focused on. These phenomena often lead to different infection rates and thereby a different number of cases. Herd immunity is a factor observed in literature on infectious diseases and so cannot be ignored. There are many different approaches of integrating this effect [1, 2], but the problem that the real epidemic is not represented by the Markovian model structure continues.
- The lack of studies especially concerning serotype shifting is a big problem and therefore it is questionable if it can be modeled in an appropriate way to estimate the effectiveness of vaccines or alternative strategies using Markovian models.

Otherwise Markovian models are necessary in cases when no further information on the epidemiology of

an infectious disease is evaluable and enough data regarding cases is given. Although especially in economic output research the implementation and the possibility to calculate the costs per case and cost per life year gained quite easily, is an important feature of this method.

Another important aspect is that this approach is often used by health economists and therefore it is necessary to implement equivalent structures for comparison with results from literature. Nevertheless this should not be the reason for choosing this approach or only this approach. In several cases parallel development of different modeling techniques for one simulation question is necessary.

3 ODE models

Another top – down method in modeling infectious diseases are ordinary differential equations (ODEs). This dynamic method is based on the concept of Kermack and McKendrick and uses the well defined ODE theory.

ODEs use population groups in contrast to cohorts. The basic equations for an SIR epidemic are defined as follows:

$$\dot{dS} = -\beta IS$$
$$\dot{dI} = \beta IS - \mu I$$
$$\dot{dR} = \mu I$$

As can be seen the dynamics of the infectious class depends on the following ratio:

$$R_0 = \frac{\beta}{\mu}$$

the so-called basic reproduction number (also called basic reproduction ratio). This ratio is derived as the expected number of new infections from a single infection in a population where all subjects are susceptible. Fig. 2 shows a classical solution of an SIR basic model with only one homogeneous population.

To implement different, for example age dependent, parameters the model structure has to be implemented the same way in parallel more often. This expands the parameter list and the complexity regarding the solution methods. Concurrence of different serotypes or strains can be modeled using more complex assumptions regarding parallel incidence of illnesses and carrying rates.

Each additional feature focused on leads to a change in the equation structure and therefore to an overall model change. These changes or refinements are in general time consuming and often not communicable in interdisciplinary groups.

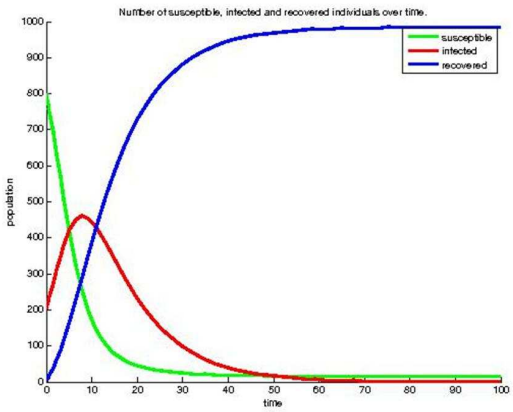

Fig. 2 Classical solution of the Kermak and McKendrick ODE – model for a saisonal epidemic. The blue graph represents the number of recovered persons, the red one represents the infected and the green the health people.

Still the main problem of ODE based concepts in modeling the spread and influence of infectious diseases is the parameterization and validation of the model parameters.

In comparison with Markovian methods ODE approaches lead to higher flexibility and better implementation of dynamic non linear effects. Both of them have strong restriction in parameterization and do not represent the real epidemiology. Therefore another concept is necessary.

Modeling infectious diseases with ordinary differential equations in interdisciplinary groups is often problematic. That is why the model description is often done using System Dynamics [3].

The model development is defined in seven steps, starting with a textual description of the problem, defining of causal loops and the conversion into flow diagrams. In most cases, if the epidemiological structure of an infectious disease is known, it is quite easy to define the causal loops.

Afterwards the implementation and first simulation of the concept has to be done. In this step the problematic of parameterization of the up to now qualitative approach with quantitative values occur. Outlining alternative scenarios, refinement of the feedback loops represented by causal loops and discussion as well as implementation of the changes are the last steps of the model development process using System Dynamics.

The main benefits of ODE modeling of infectious diseases in comparison to Markovian models are the higher flexibility, the chance to model the influence of different population groups and their influence on the overall behavior as well as the well known theory regarding stability analysis [4]. Furthermore effects like serotype replacement can be modeled, however the problem of parameterization is not solved

The lack of studies in literature concerning serotype shifting is a big problem and therefore it is questionable if it can be modelled in an appropriate way to estimate the cost-effectiveness of vaccines. This problem leads to a significantly different approach for modeling of infectious diseases and a change in the perspective from top – down to bottom – up.

4 Agent based models

Agent based models are bottom – up models of a heterogeneous population of agents and their social interactions. The summed results of the micro level interactions can be used to calculate the interesting macro level behavior.

Because of building up each single person of interest using a simplified computer model, the so defined modeling structure is part of the bottom – up methods. In contrast to the other presented strategies the real life implementation of the spread of epidemics is in center of interest.

Social contact modeling is a key element in simulating infectious diseases because a transmission can only happen when people meet each other. Therefore additional knowledge on the demographic structure, the social interaction rates, etc. is important.

Based on these definition and ideas the following modeling structure for simulation of epidemics with agent based techniques can be defined:

The agent based modeling structure can be set up in four main parts: The population part, the social part, the infection part and the economic evaluations. The agents in this model are single persons with attributes age, gender and infection state. In the population part the population is calculated, people get older and die and babies are born using real life data for Austria. For the spread of an epidemic people have to meet each other.

In the social part contact lists of each single person can be defined. These connections are varied every time step by resolving some old and defining new connections using different algorithms for evaluation of social structure behaviour over time. An alternative approach is modeling the social structure in more depth by defining following separate parts where social contacts occur:

- households,
- schools,
- working places,
- spare time, and
- places of special interest (hospitals, foster homes…).

Using this advanced structure additional scenarios like the effect of closing schools during an epidemic or the reduction of contact rates of interest by using the assumption, that a part of the population wears breathing protection can be calculated.

In the infection part a person that stays in contact to an infected person is tainted with an individual probability depending on the own age, the serotype coming in contact with, the underlying diseases of the person.

The fourth and last part includes the economic evaluations, the connection between the cases of illness and the age dependent course of disease and thereby the associated costs per case.

Using this modular structure the over all concept as depicted in Fig. 3 can be implemented.

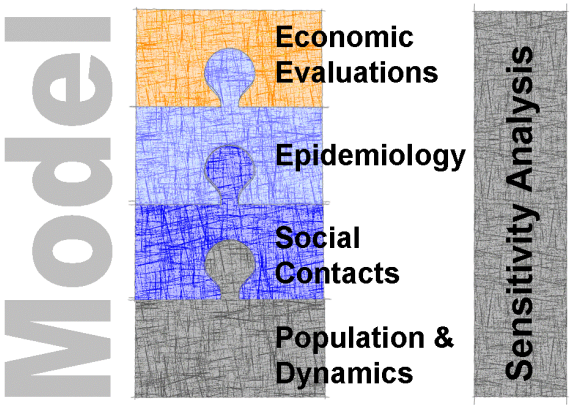

Fig. 3 Definition of the underlying modular structure for agent based modeling of strategies for infectious diseases. Sensitivity analysis for the main influence parameters and for evaluation of the stability has to be performed.

Because of these additional influence factors taken into account and because of the possibility of refinements regarding sex, age, income, and so on the number of parameters increases rapidly. This seems to be a problem that cannot be handled in the first moment. The main difference in comparison with Markovian models and ODEs is that the used parameters are not that complex and especially demographic and contact parameters can be identified.

Furthermore the clear structure offers cooperation in interdisciplinary groups.

Using single parameter sensitivity analysis, the main influence parameters can be identified. A multi parameter sensitivity analysis is used for evaluation of the overall system behavior and for stability analysis. As this modeling approach is not based on an equivalent stability theory like ordinary differential equations or Markov chains, mostly Monte Carlo techniques are used.

As agent based simulation deals with a high number of single persons parallelization of computation processes has to be performed, otherwise simulation time runs out of time.

5 Results

Current models work with reduced macroscopic or static models and interpretations. Classical methods of dynamic modelling via a system of ordinary differential equations are in this case not conducive, as infection probability varies greatly with age and furthermore many studies take into consideration a very long time span, making demographic shifts of the basis population an important influence. Furthermore, contact probabilities between persons of different age groups have to be taken into consideration and included in the model parameters. With the necessary division of age groups into years of birth added, the resulting least possible number of parameters is growing and growing. Hence the according differential equation models could not be parameterized, making control in the context of an interdisciplinary model development impossible. The aim is the development of dynamic models that are both calculable and identifiable.

The modeling and simulation concept of agent based systems to calculate scenarios on the spread of infectious diseases is the most flexible strategy presented. The main benefits are:

- the wide range of scenarios that can be evaluated,
- that herd immunity and serotype/strain replacement are generated by the defined contact structure and are no fixed parameters or complex numbers, and
- the readability of all model parts also for non mathematicians. This leads to higher reliability of the model.

The main pitfalls of the agent based modeling concepts are the longer development and testing time, as well as an extended computation time. These two problems have to be handled by using reusable sub models for social interaction, demographic change, etc. on the one hand and parallelization of the computation on the other hand.

Combination of the presented modeling methods can lead to faster development time giving the same quality of results. Using different approaches at the same simulation run might be useful when some of the model parts do not have to be modeled in detail or if reliable data for the macroscopic system is given.

Results can be implemented in systems for calculating new strategies for vaccination programs. Current work considers two or more concurrent serotypes as herd immunity and serotype replacement affect each other and become indistinct.

In cases where only little knowledge about the epidemic of an infectious disease is given, Markovian models or even a single Decision Tree can be used to get first results. But it has to be taken into account in every case, that with this method no real epidemic behavior can be modeled dynamically.

6 Outlook

Further work in model behavior and flexibility testing will be done. Assumptions regarding combination of different modeling techniques, especially ordinary differential equations and agent based concepts, are part of the current work.

The actual problem solution structure of questions in health technology assessment (HTA) is shown in Fig. 4. In most cases a method is chosen before closer insight in the data structure is available. In most cases no evaluation of the best solution method is performed and no feedback regarding data quality and/or data structure is given.

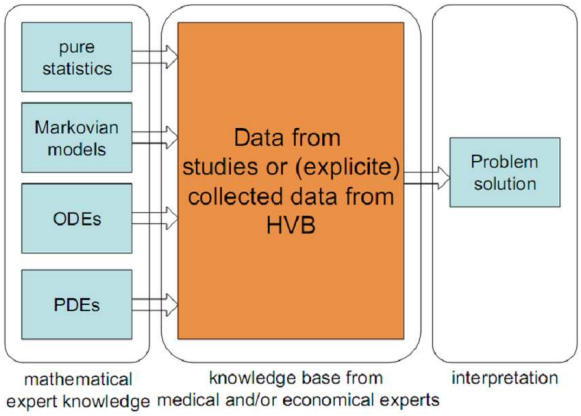

Fig. 4 State of the art problem solution in HTA and modeling of questions regarding the spread and dynamics of infectious diseases.

In the current work, especially parallelization and comparison to classical implementation strategies of agent based models are in the focus of interest.

Parallel concepts for better combination and communication between problem definitions/modeling questions, data resources/data specialists and modeling experts are tried to be realized. In Fig. 5 an idealized concept for complex problem solving in the field of infectious disease strategy evaluation is shown. The main additional points in this flow diagram are the model decomposition and the green marked feedback arrows, which guarantee a better problem solution strategy for future questions.

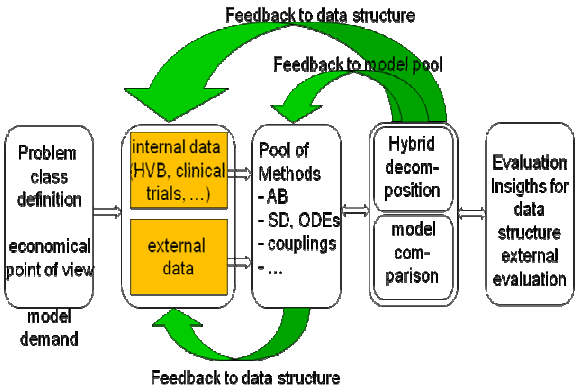

Fig. 5 Flow diagram with feedback loops and model decomposition for the overall problem solving strategy for infectious disease questions.

In comparison with the chart shown in Fig. 4, Fig. 5 starts with the general question of interest and thereby defines the model demand. The modeling concept for an infectious disease depends on the general modeling question as well as on the point of view and the given data structure. This is one of the main differences in comparison with the state of the art structure: The given data structures have direct influence on the modeling method that is chosen for problem solving.

Based on the chosen method feedback on the data sets will be given during implementation and scenario calculation.

Furthermore a combination of different modelling approaches from the fields of cellular automata or agent based model theory, discrete event simulation as well as, e.g., System Dynamics or flow models on local scales will be employed. An important part of the ongoing work is the development of applicable and reusable part models (social model on agent basis, contact modelling at workplaces/schools with cellular automata structures, ...) and the definition and implementation of applicable interfaces.

In order to achieve international HTA expert recognition for these new model approaches and to allow for comparison with "state-of-the-art" methods, current standard models will be developed in parallel (in this field Markov basic models) with a following mathematical treatment and discussion of the limits of non-implementable part areas. This will make it possible to analyze questions such as the cost-benefit ratio, which have so far not been able to be sufficiently examined taking into consideration different aspects of application, such as herd immunity, etc.

An important advantage of this approach is the examination of the serotypes contained in the vaccine in the context of rival strains. Only this will make it possible to detect and classify the effect of serotype shifts in a dynamic model.

The cost-benefit ratios that will be arrived provide an independent basis for evaluation for the decision-makers.

7 Goal

Developing new standards of evaluation of therapeutic and preventative measures via models and simulations in the area of infectious diseases, based on existing interdisciplinary methods in health economics is a main goal of the current work. The results will lead to a structural change in the working methods for health technology assessment (HTA) and allow the use of cost-effective dynamical models to supplement expensive and protracted empirical studies, thus improving the ability to predict reliability and precision of medical treatment, interventions and preventative measures.

Advanced modelling and simulation techniques as well as coupled simulation techniques should be used as a decision support system for vaccination strategy evaluation and as standard method in the planning phase of national pandemic plans.

8 References

[1] GP. Garnett. Role of herd immunity in determining the role of vaccines against sexually tramsmitted disease, *The Journal of infectious disease*, 2005 Feb 1;191 Suppl 1:S97-106

[2] AR. McLean and SM. Blower. Imperfect vaccines and herd immunity to HIV, *Proceedings. Biological sciences / The Royal Society*, 1993, 253(1336):9-13.

[3] H. Bossel. Systeme, Dynamik, Simulation: Modellbildung, Analyse und Simulation komplexer Systeme, *Norderstedt : Books on Demand*, 2004, ISBN 3833409843

[4] G. Endel, I. Schiller-Frühwirth, N. Popper, G. Zauner, F. Breitenecker. Comparison of Different Static and Dynamic Simulation Techniques for the Influence of Children Pneumococcal Vaccination, *Value in Health / Wiley Blackwell*, Vol. 12 no. 7 (2009), ISSN: 1098-3015.