Modellierung im Gesundheitswesen

Tagungsband des Workshops im Rahmen der Modellierung 2014
Modeling and Simulation of Screening: Methods, Technologies and Applications in the Austrian Setting

Günther Zauner, Christoph Urach, Patrick Einzinger, Niki Popper, Irmgard Schiller-Frühwirth, Gottfried Endel

Abstract

Modeling techniques and the decision process for research question definition using PICO (Population, Intervention, Control/Comperator, Outcome) in HTA (Health Technology Assessment) problems are described briefly in the first part of the paper. The second part deals with some aspects of realized screening questions regarding cost effectiveness of organized breast cancer screening in Austria and a screening program for abdominal aortic aneurysm. In case of breast cancer a closer look on model calibration for natural history behavior of tumor growth/development is given.

Keywords: screening, agent based modeling, microsimulation, modeling methods, abdominal aortic aneurysm, breast cancer screening.

1 Introduction

In modern health security systems the transition between health care and treatment towards prevention is happening stepwise. The different strategies like healthier living regarding food, sports and reducing overweight as social factors go together with services of the health care providers in the field of medical check-up and screening. Thereby chronic illnesses as well as cancer are of main interest because they affect a lot of people and looking at the billing they cause a main proportion of costs used for care. Due to these reasons screening gets more and more in focus. Quality assessment and strategy planning is an important longtime strategy to guarantee best treatment and thus get a high participation rate regarding the target population as well as best cost effectiveness,. The health technology assessment (HTA) (Busse and Perleth 2008) of the screening strategies can follow the same process as well known techniques in HTAs of vaccination strategies or HTAs for drug evaluations (Zauner et al. 2010b; Zechmeister et al. 2009). Questions to be answered occur in health policy making when

1. a new screening technology comes into discussion or
2. the providing strategy is changed, for example when an opportunistic screening is changed to an organized screening with quality assurance, or
3. bounded resources have to be distributed to a small section of population
the evaluation process has to be started using modeling and simulation to provide a broadly accepted, objective framework showing the uncertainties and providing scenario calculation as well as sensitivity analysis. To satisfy these requirements a well structured process is necessary and different modeling methods have to be taken into selection process.

2 Research Question

As connective link between the raising health policy discussion of a new screening program and the implementation of the technology assessed, the definition of the main questions answered for HTA have to be defined. Hereby national or international guidelines and law have to be taken into account. In structured work this is often done using the PICO (Straus et al. 2005) framework. PICO describes a standardized question including:

- Population,
- Intervention,
- Control,
- Outcome.

This paper does address the issue of how to generate a reasonable question for the modeling process on basis of the PICO question. The processes of selecting a question for decision-making (prioritization) and of evaluating the findings within a broader political context (appraisal) are not covered in detail in this paper. These aspects are, particularly in Austria, largely influenced by political decisions. Nevertheless different modeling techniques are in the focus. They provide the implementation of different aspects and overall view on the problem to be answered. Modeling and simulation act as decision support tool to:

- calculating the basic scenario representing in most cases the Control,
- realization of the Intervention set up,
- calculation of scenarios,
- transformation of the simulated results to the defined output format and representation of the Outcome.

All work should be done based on the defined Population as given by the PICO research question.

2.1 Definition

Working on a problem in HTA usually at first requires a clarification of the potential decisions, the definition of the population, the intervention, the comparator and the outcomes of interest, the
PICO question. In this phase the question has to be worded as precisely as possible, while at the same time the feasibility and necessity of modeling have to be discussed. Knowledge of the health care system is essential to understanding the various paths the HTA process can take. The PICO question itself is the starting point - the first point where the potential quality of an assessment is decided. It should be defined using an interdisciplinary group to end up in:

- A well defined research question,
- A description that fits in detail the decision demand of politics, and
- A specification for the model selection and parameterization process.

Having arrived at the PICO question, it needs to be identified which part of the HTA process can be supported by modeling and simulation. It has to be decided which research questions can be answered using models and what steps have to be taken in order to be able to do so. These decisions should be taken by a steering group covering the various aspects that come into play.

2.2 Data

Based on this, the problem’s data and structure have to be analyzed. The raw structure of data gives additional information about knowledge depth and real world data availability for parameterization of the following model and decision process. The model structure decision process can be realized in a feedback structure as shown in (Zauner et al. 2010a). In Austrian and international setting the selection of the target population becomes more and more important. For instance, based on the guidelines of Vienna health promotion institutions it is stated out that: especially different health care needs of men and women, based on biological sex and social gender are factored in gender mainstreaming. The gender mainstreaming aspects have to be discussed in each direct and indirect action of health promotion. In the next section potential modeling methods and the modeling technique selection process, which represent an iterative process performed within the steering group, are described.

3 Model structure decision process

The model structure decision process for screening models is influenced by a broad variety of demands and guidelines. Thus it is necessary to choose a modeling method that is not only capable to fit the problem adequate, but also easily comprehensible to guarantee that it can be discussed in an interdisciplinary team. Furthermore it should offer a structure that can easily be validated by the external decision makers. This point is essential in getting an objective decision process, and is supported by following quotes:

"Complexity means distracted effort. Simplicity means focused effort." (De Bono, 1998)
"Make everything as simple as possible, but not simpler." (Albert Einstein)
The following interpretation can be given regarding modeling and simulation in screening settings: it tries to provide the best possible description of a system in the simplest possible way. This avoids producing descriptions that are over-full as well as descriptions that fail to explain enough. To get the insight in the complexity and to choose an adequate model used for modeling a screening program preliminary work has to be done. Especially:

- Defining the detailed research question
- Setting up an interdisciplinary working team and
- Clearly defining the result transfer to the decision makers (integrating the decision process into the question setup or at least find a consensus of what are the main points of interest answered during project phase).

### 3.1 Characterization and model types

As a model is always a simplified representation of reality different possibilities to implement changes in state variables can be realized. When and how this happens is part of the modeling technique. These are in general:

- Discrete time steps
- Continuous time flow
- Discrete events.

Some of these concepts have to be taken because they originate directly from the used technique. Others are chosen because of the model concept, investigated question or simplicity. This can be the reason if there is not much improvement by choosing a more complex method.

Classification can be done using different strategies: White-Box versus Black-Box modeling describes the kind of information and functional coherences used. Top-Down versus Bottom-Up provides a concept of the view on the population. Another classification of population can be cohorts versus microsimulation versus agents (Weinstein 2006). Continuous / discrete time and discrete events characterize the simulation progress (Bossel 1992; Cross and Moscardini 1985).

**Decision trees** are tree-like hierarchical ordered graphs or models representing options for actions (strategies) and sequences of possible events following these actions. For each potential strategy, there are paths that represent all possible sequences of events. The event probabilities can vary with the chosen medical action or with patient characteristics. At the end of the decision tree, each path leads to a certain outcome value (for a quite simple example see Schöffski and Schulenburg 2008). The main advantage of such a modeling system is the clear structuring of decisions and consequences in a stochastic decision problem. The disadvantages are substantiated by the fixed time horizon and by the fact that the decision tree can become unwieldy for recurring events.
Markov models describe the progression of a disease through a number of disparate states. Patients or cohorts are always in one of these predefined states. All events of interest are modeled as transition from one state to another. Updates of states are performed at discrete time steps. Transitions are usually integrated into the branches of decision trees that structure the decision problem. Markov models are based on the Markov assumption: The future of the system depends on the present state only and not on the history.

Markov cohort models are typical representatives of top-down model approaches. Furthermore they provide a relatively simple handling, so they are preferred in many areas of HTA (Briggs and Sculpher 1998; Sonnenberg and Beck 1993). Advantages are a direct projection of data which allows relatively easy parameterization, transparent calculation that is also understandable for non-modelers, and that there is no need for computer programming; it can also be calculated in programs like Microsoft EXCEL. Disadvantages mainly occur due to the reason that dynamic effects that obviously exist in real systems are completely ignored. Furthermore modeling of feedback loops is not possible. A Markov model is not a representation of a whole adaptive system over a time but only follows a single person or cohort over time, disregarding the rest of the population. So they are restricted in their level of detail.

Ordinary Differential Equations (ODEs) occur in HTA questions in screening only indirectly, namely, using System Dynamics (SD). SD is an approach for modeling dynamic systems which evolve continuously over time. It depicts systems consisting of a set of levels (also called stocks; one can imagine them as reservoirs or water tanks) which are connected by flows or rates (one might imagine flows as pipes). Each level stores a numerical state. Flows change the value of states over time (an example of SD in breast cancer screening: Tejada et al. 2013). As SD is a top-down approach which can handle modeling interaction: the flow rates of the levels (or cohorts) can depend on the levels themselves and also on other levels. This type of interaction would not be possible in Decision Trees or Markov Models. Nevertheless there are a few situations when a SD approach should not be the chosen (Brennan et al. 2006):

- If the population is inhomogeneous in multiple attributes which are important for the problem, as models of a large number of cohorts can become quickly difficult to handle.
- If the (inter-)actions of individuals should be modeled.
- If the population is small and the assumption that the number of individuals in a level can take on continuous quantities would lead to wrong results.

Discrete event simulation (DES) is one way of building up models to observe the time based (or dynamic) behavior of a system. There are formal methods for building simulation models and ensuring that they are credible. During the experimental phase the models are executed (run over time) in order to generate results. The results can then be used to provide insight into a system and as a basis to make decisions. The main characteristic of Discrete Event Simulation is that the system state does only change at certain time points when events occur. Time moves from one of these events to the next, the time in between is of no relevance because no event occurs.
DES in general allows a very clear differentiation between structure and dynamic behavior of a system which is one of the main benefits. Reducing this dynamic behavior to a series of events that take place at certain points in time makes this method of modeling a very useable approach for a wide variety of applications. The main problem with event based modeling is the occurrence of events that take place at the same point of time. This may quickly lead to a distortion of the course of events, based on the wrong order of changes to the system.

**Agent based modeling (ABM)** is a relatively young discipline that has become possible with powerful computers. The main goals of ABM are:

- The possibility to model details satisfyingly exact.
- Creation of dynamic effects that cannot be created with other models and that cannot be represented by parameters.
- Models with natural and descriptive structure and data.

ABM provides a natural description of the real system. That means that agents in the models look like components in reality and act in a way that can be observed in real systems. Parameters in the model correspond with quantities that occur in reality. The technical advantages are the possibility to model dynamic effects by well-known rules, due to flexible model environments that provide scalability for size and details in a simple way. The advantages in application are mainly that no knowledge about mathematical theory is necessary, and model structure is very clear and therefore the interdisciplinary partners in a screening HTA can give their input without having advanced knowledge in computer techniques. The main disadvantages are based on the fact that small errors in the description of single agent behavior can influence the overall system reaction. Besides that there are only a few mathematical methods for analyzing ABM. Often statistical evaluation of results is the only possibility.

### 3.2 Model decision process

Summing up, there is a list of modeling methods fitting to HTA questions in screening applications on different level (point of view, details) and in different quality. To get best results and to follow best practice it is necessary to realize the modeling method decision after:

- The definition of the main research question and
- Basic evaluation of given data structures for parameterization.

Another often postulated demand, especially from decision maker’s point of view, is that the model under discussion has to be comparable with selected published international work. That influences the model structure selection process.
4 Example applications for the Austrian setting

Two substantially different examples for model based HTA in screening questions for the Austrian setting realized by the authors are presented in the following section. For this purpose parts of the whole work are explained and discussed briefly. The first example deals with a chronic disease - abdominal aortic aneurysm (AAA) - and the other one deals with organized breast cancer screening for women, the most frequent cancer in women in developed countries.

4.1 Abdominal Aortic Aneurysms

Abdominal Aortic Aneurysms (AAA) is a disease which describes an overlarge abdominal aorta, which may lead to its rupture which usually ends fatal. AAAs concern about 2 - 6% of all men aged 65+. Prevalence for women is only about one fifth of that of men, but the probability of ruptures is about four times higher. Ultrasonic screening is used for diagnosis with about 100% sensitivity and specificity, it is cheap and therefore considerations about addition of organized AAA screening to the medical check-up program in Austria are obvious. The PICO research question is defined as follows:

- **Population**: 65-year old people over 20 years
- **Intervention**: organized Screening of the age group, including follow up strategies
- **Control**: no screening strategy
- **Outcome**: number of ruptures, fatality rate and cost effectiveness ratio

Boundary conditions regarding modeling and simulation: The simulation model should allow the evaluation of AAA screening programs, especially testing assumptions which cannot be examined in real life experiments or clinical trials due to ethical, technical (time horizon) or cost-related reasons. The modeling research question is therefore to

- develop a simulation model for development, screening, treatment and corresponding costs of AAA in Austria
- analyze AAA-induced cases for 65-year old people over 20 years
- simulate the influence of key factors for AAA development and rupture
- assess organized screening following EUnetHTA core information (Kristensen et al. 2009)

Parameterization of the model is mainly performed based on information created within the IFEDH (FFG grant number 827347) project. This information was supplemented with reimbursement data as well as results from studies. Aggregation of the collected data to technical parameters is described in (Abdominal Aorta Aneurysm – Cost-Effectiveness Analysis on Introduction of Organized Screening in Comparison to Current Practice in Austria), assumptions for screening strategies or fuzzy parameters were also realized within IFEDH. One basic point in setting up
parameters was the comparability (only on parameter level, not the simulation technique) with work done by EUnetHTA.

Within the IFEDH-project a model structure for AAA screening was developed. Thereby using an agent-based model was suggested consisting of four interacting parts (see Figure 1): Population module, Disease progression module, Treatment module, Protocol module;

The model structure decision process was realized by people from Main Association of Austrian Health Security Institutions (having the role of the decision maker), HTA-experts, statisticians and experts in the field of dynamic modeling and simulation from dwh simulation services and Vienna University of Technology. The ABM structure was used due to the reason that not all scenarios where defined in the beginning of the process and therefore highest flexibility was one of the key decision arguments. Beyond that weaknesses in realized international models have been detected and discussed. As third reason for using modular ABM structures the re-usability of well defined model parts was in focus.

Fig. 1: Overview of the modular modeling structure based on formular point of view

The population module initializes the targeted population and simulates its progression over the modeled time horizon, for example aging or non-AAA-related death. Disease progression, in this case the growth of the abdominal aorta, is implemented in the disease progression module (see Figure 2 for representation of disease classification). Surgeries belong to the treatment module while the protocol module is responsible for the recording of values of interest.

The concept behind this structure lies in flexibility and reusability. Disease or treatment specific assumptions only need to be changed within the corresponding modules. The protocol serves two purposes. It records values and performs analyses, for example the calculation of costs of life years gained. The base run simulates the development and treatment of 65-year old people (in 2012) over 20 years with and without organized screening. Screening strategies are compared, the greatest possible impact is shown by (hardly realistic) 100% screening participation. Because of different epidemiological behaviour, the population is evaluated separately for men and women. Afterwards, sensitivity analysis is performed to assess impreciseness of the results caused by fuzzy parameters.
From modeling theory point of view it can be stated that in the Austrian context agent based modeling techniques are identified as most flexible and well suited for modeling the AAA incidence as well as screening and treatment.

4.2 Cost effectiveness analysis of breast cancer screening in Austria

Breast cancer is the most common cancer in the female population in Austria. Mammography screening trials have been shown to reduce breast cancer mortality by 15% (Gøtzsche and Nielsen 2009) to 25% (Vainio and Bianchini 2002) in screened women who are more than 50 years of age.

The aim of this study was to evaluate the cost-effectiveness of an organized screening program for the early detection of breast cancer compared with an established opportunistic screening and to identify factors influencing the clinical and economic outcomes from the Austrian health care system perspective. The PICO research question is defined as follows:

- Population: asymptomatic women, 45 to 70, at normal risk for breast cancer
- Intervention: Organized Screening for breast cancer with mammography
- Control: Opportunistic Screening for breast cancer with mammography
- Outcome: Death cases avoided, incremental cost effectiveness ratio

A decision analytic microsimulation model was applied as the appropriate modeling approach. The microsimulation can be used to handle all necessary requests of the modeling structure. Due to the reason that no influence in behavior (like the probability to take part for each single woman)
is under discussion it is not called for the flexibility of an ABM. The Markov model consists of 3 main components:

- comprising a breast cancer natural progression pathway (Figure 4), including clinical diagnosis;
- an opportunistic screening and
- organized screening pathway.

The health states in the Markov Model were defined as no cancer, undiagnosed and diagnosed ductal carcinoma in situ (TIS), undiagnosed and diagnosed invasive local, regional, distant cancer, death from breast cancer and non-specific mortality attributable to non-breast cancer causes. The structure of the model is similar to other models already used in the evaluation of breast cancer screening programs. This guarantees the comparability. The majority of economic models for breast cancer screening in the literature have classified cancers by tumor size like MISCAN (Gelder et al. 2009). However, Statistic Austria data do not categorize cancer detection rates by tumor size, but
data are reported as EOD (extent of disease coding), a simplified TNM (Tumor, Node, Metastasis) classification.

A model that evaluates the cost-effectiveness of the current opportunistic screening scenario in Austria compared with an organized screening program is developed using TreeAge Pro 2013 (TreeAge Software Inc) software.

As an example of cooperation of different disciplines in one research project leaded by a HTA-expert and for more detailed explanation of key assumptions the calibration (Taylor et al. 2007) of the breast cancer incidence (not detected cases) is chosen: Incidence and the prevalence of mamma carcinoma in 40-year-old women are important parameters of the model. While there are good
data on breast cancer diagnoses in Austria, it is naturally very difficult to estimate these values. Therefore these parameters were calibrated with data on diagnoses.

The number of diagnoses with breast cancer of women at age 40 and above during the years 1987–1991 from the Austrian cancer registry were included. This time was prior to the introduction of opportunistic cancer screening in Austria. Therefore it was possible to use the model without the implementation of screening. For every five-year age group this resulted in the numbers of cases per stage. They were then divided by the number of person-years of women in the respective age group and multiplied with 100000 to get the number of breast cancer diagnoses per 100000 women.

For calibration purposes the model of the natural history of the disease and the clinical diagnosis was re-implemented in the simulator AnyLogic, because TreeAge Pro has no automatic calibration capability and it is also harder to compare the output of simulation runs over time with data. AnyLogic uses the OptQuest optimization engine for calibration purposes.

Figure 6 shows a graphical representation of the process that each woman goes through. It is the same as in the original model implementation. The model uses 38 parameters, of which 28 were calibrated. The calibration experiment reads the desired diagnosis numbers (a total of 48 data values for TIS, local, regional and distant cancer in twelve age groups) and compares them with the simulation results. For this purpose, at the middle of each five-year interval it counts the women that are alive and undiagnosed. This number multiplied with the observation period of five years gives the number of person-years. Then after the interval it divides the number of diagnosis that occurred during the interval by the number of person-years and multiplies it with 100000 in order to get a value comparable to the data.

The objective function of the calibration was composed as follows: For each cancer stage (TIS, local, regional and distant) the square root of the average of the squared differences between all ten data values and simulated values was calculated. Additionally, as 40% of invasive cancers should be preceded by TIS, the absolute difference of the simulated proportion and 0.4 was also calculated. All these five single values were weighted with 100 divided by the mean of their data values in order to get comparable scales.

The calibration procedure started with 500 automatic iterations of the simulation. Every iteration consisted of at least two and a maximum of 100 replications with identical parameters to account for stochastic variance. The ranges for the calibration parameters were set rather broad, from 0 to 0.001. After this automatic calibration parameter ranges were identified and therefore the basis for improving the calibration by manually changing parameter values and comparing the results was given. The fit for TIS and regional cancer diagnoses is acceptable. For distant diagnoses the simulation results in moderately higher values in younger age and lower values in older age. The proportion of invasive cancers preceded by TIS is with 0.37 a bit lower than the target proportion of 0.4.

The gathered results of model calibration give an adequate starting point for the next steps realized for the mammography screening project based on the defined PICO question. In this example the necessity of interaction between the specialists in different domains (medicine, data
Fig. 5: Implementation of the women in the model. The state chart on the left shows the possible progression through stages. Each woman starts in an undiagnosed stage according to the start prevalence. Progressions from healthy to TIS and to invasive cancer, or from TIS or an invasive cancer stage to the next higher stage are possible. Additionally, a woman with TIS may become healthy again, and TIS as well as invasive cancer stages can get diagnosed. Death is also possible in each undiagnosed stage. The action chart on the right shows the sequence in each time step.

knowledge and modeling) was very important due to lack of real world high evidence data on natural history of disease progression.

5 Results and Outlook

The literature and theory work done by the working groups of the paper in the field of HTA, modelling techniques in health economics, as well as developments in hybrid model coupling and modular modelling led to question solving ability for screening issues.

Due to rising of more complex HTA questions in future, the solution techniques need to be improved periodically. The interdisciplinary model decision process based on PICO and given data structures/boundaries seems to become more important in future; simultaneous the time span for
question solution is lowering. This is why modular modeling concepts and a broad knowledge on different techniques are in focus. Standardized or automated parameterization of model parts are under development and are promising regarding speeding up decision support with necessary actuality and quality of parameterization.

Conflict of interest and acknowledgement

dwh and our content experts wish to disclose they have no financial interests or other relationships with the manufacturers of commercial products, suppliers of commercial services, or commercial supporters. The IFEDH research project was funded by the COIN - Cooperation and Networks program of the Austrian Research Promotion Agency (FFG), the national funding institution for applied research and development in Austria. The project is headed by dwh and runs from October 2010 to September 2012.

References


European network for Health Technology Assessment, EUnetHTA. In: International journal of technology assessment in health care 25(S2), pp. 1–8


Straus S. E., Richardson W. S., Glasziou P., Haynes R. B. (2005) Evidence-based medicine: how to practice and teach EBM. In:


Zauner G, Miksch F, Popper N, Endel G, Schiller-Frühwirth I, Breitenecker F (2010b) PIH47 LONG-TERM EFFECTS OF CHILDREN PNEUMOCOCCUS VACCINATION: AN AGENT BASED APPROACH. In: Value in Health 13(7), A383–A384