Dealing with uncertainty in health economic decision modeling

Applying statistical and data science methods

Inauguraldissertation

zur Erlangung des akademisches Grades

Doctor of Public Health

der Fakultät für Gesundheitswissenschaften der Universität Bielefeld

vorgelegt von
Stefan Scholz
M.Sc. in Public Health
M.Sc. in Statistischen Wissenschaften

1. Gutachter: Prof. Dr. Wolfgang Greiner
2. Gutachter: Prof. Dr. Ralf E. Ulrich

Berlin, 2020

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ABM	Agent-Based Model
AIC	Akaike Information Criterion
API	Application Program Interface
BIC	Bayesian Information Criterion
BPD	Broncho-Pulmonary Dysplasia
CDF	Cumulative Distribution Function
CEAC	Cost-Effectiveness Acceptability Curve
CEAF	Cost-Effectiveness Acceptability Frontier
CI	Confidence Interval
DES	Discrete Event Simulation
DMARD	Disease-Modifying Antirheumatic Drug
DSA	Deterministic Sensitivity Analysis
EHMD	Exclusive Human Milk Diet
EVPI	Expected Value of Perfect Information
EVPOT	Expected Vaccine Protection Over Time
GAM	Generalized Additive Model
GDP	Gross Domestic Product
GLM	Generalized Linear Model
HEDM	Health Economic Decision Modeling
HIV	Human Immunodeficiency Virus
HTA	Health Technology Assessment
ICER	Incremental Cost-Effectiveness Ratio
IMD	Invasive Meningococcal Disease
IPD	Invasive Pneumococcal Disease
ISM	Individual Sampling Model
LYG	Life-Year Gained
MSM	Men-who-have-Sex-with-Men
MSW	Men-who-have-Sex-with-Women-only
NB	Net Benefit

NBPP	Non-Bacteremic Pneumococcal Pneumonia
NEC	Necrotizing Enterocolitis
NHB	Net Health Benefit
NMB	Net Monetary Benefit
ODE	Ordinary Differential Equation
OLS	Ordinary Least Squares
OTA	U.S. Office of Technology Assessment
PCV	Pneumococcal Conjugate Vaccine
PPSV	Pneumococcal Poly-Saccharide Vaccine
PrEP	Pre-Exposure Prophylaxis
PSA	Probabilistic Sensitivity Analysis
RA	Rheumatoid Arthritis
SE	Standard Error
SHI	Statutory Health Insurance
STD	Sexually Transmitted Disease
STI	Sexually Transmitted Infection
WHO	World Health Organization
WTP	Willingness-To-Pay

ABSTRACT

Health economic decision modeling is a widely used method to support decision makers in the health care sector choosing cost-effective interventions. Modeling allows to combine evidence from various sources and to compare different treatment strategies that exceed the feasible number of comparators in clinical trials. New, increasingly complex health technologies need to be reflected by more complex models and need to be informed by more data. Additionally, more detailed models come with more and new types of uncertainties surrounding model input, model structure and methodological choices.

The aim of the present dissertation thesis is twofold: First, statistical and data science methods are explored to deal with the increased data demand of more complex models and to be applied for the classical forms of uncertainty. The second goal is the exploration of a new source of uncertainty originating from the heuristic nature of new model types like agent-based modeling – namely uncertainty associated with the algorithms used to simulate agent behavior.

The thesis is comprised of eight articles published in international, peer-reviewed journals. The first six articles are dedicated to the application of statistical and data science methods to model uncertainties regarding parameters, methodological and structural choices. The publications show how new methods, for example web-scraping, can be used to inform behavioral, epidemiological as well as health economic input parameters of models. They also present how uncertainties can be explored and communicated to decision makers. The last two articles explore the influence of algorithmic uncertainty in agent-based models used in infectious disease modeling. The results show the extent in which matching algorithms influence the spread of a sexually transmitted infection.

The thesis demonstrates that new statistical and data science methods allow the use of more detailed, complex models to answer very specific research questions. This allows decision makers to find tailored, cost-effective interventions. However, these new model types also come with new uncertainties that must be communicated to the decision makers to fully account for the uncertainty surrounding the model results.

CHAPTER 1

Modeling in Health Technology Assessment

1.1 Introduction

The World Health Organization (WHO) defines the highest attainable standard of health as a basic human right. Subsequently, every human being should be able to access the health care services they need to achieve this health state [1]. While health care and good health are in themselves a desirable goal, they are also a necessity for economic growth and prosperity of countries [2, 3]. From an individual perspective, steadily increasing the private expenditure on health with increasing income seems rational [4]. For publicly financed health care, however, an optimal level of health care expenditure might exist, after which additional expenditure might only yield marginal effects on population health but detrimental effects on economic development [5, 6, 7]. Driven by demographic transition, rising incomes and new technologies, many developed countries find themselves spending more than this threshold and are consequently under increasing pressure to contain public spending on health care [8].

The most popular measure to respond to the problem of expanding health care budgets is Health Technology Assessment (HTA). The first application of a systematic technology appraisal in the health care setting was carried out by the U.S. Office of Technology Assessment (OTA) in 1976 and quickly spread across other countries [9], arriving in Germany in the 1990s [10]. The basic idea of HTA is that a new technology A – defined as a new procedure, product or structure/organization in health care provision – should be systematically

assessed with regard to it's medical (i.e., efficacy, safety), economic, ethical, organizational, socio-cultural and legal dimensions. The new intervention A should be compared against one (B) or several ($B_1, B_2, \ldots$) alternatives, representing the current technology or a placebo or "doing nothing" if no comparator exists [11]. Thus, HTA-reports ideally provide decision-makers with a comprehensive, unbiased overview of the potential impact of a new technology on the health care system compared to the current status-quo. While usually all domains should be considered in the decision, the economic domain of HTA is particularly useful restricting health care expenditure and containing budgets.

HTA relies on modeling as an essential tool to combine different sources of evidence and to explore many different scenarios. A variety of modeling techniques is available for this task. But modeling also comes with uncertainties regarding the selection of the correct modeling technique, the specific layout of the model, the estimation of input parameters and methodological choices. Not addressing these uncertainties may produce biased or wrong model results and leads to a non-optimal allocation of resources in a health care system. The aim of this thesis is to explore methods for dealing with uncertainties in health economic modeling. In the first chapter, the basic principles of economic evaluation, Health Economic Decision Modeling (HEDM) in general and the different modeling techniques will be described. The main focus of this chapter lies on the categorization of uncertainties in modeling that can be divided in *parameter uncertainty*, *methodological uncertainty* and *structural uncertainty*. The following chapters present the results from the single publications comprising the thesis and dealing with parameter uncertainty (Chapter 2), methodological and structural uncertainty (Chapter 3). Algorithmic uncertainty (Chapter 4) – as a new, emerging category of uncertainty – is described in the following chapter. The final chapter summarizes and discusses the main results of the study and concludes the thesis.

1.2 Health economic decision analysis

Research questions of health economic decision analyses can be split into three categories, regarding technical, productive or allocative efficiency [12]. Questions on *technical efficiency* are asking if the same or a higher outcome can be achieved by using less of a certain input factor. For example, a systematic review found that for patients with cystic fibrosis one daily dosis of intravenous antibiotics shows the same effectiveness as three doses per day [13]. *Productive efficiency* is also asking whether the same or a higher outcome can be achieved, but not by means of a reduction of one input factor, but by reducing one input factor at the extent of another. Switching from a surgical procedure to a medication for the same indication can be seen as an example.

Finally, *allocative efficiency* is concerned with distributing health care expenditures – i.e., the input factors – in such a way, that a general, comprehensive measure of outcome is maximized across the population or within the health care system.

Irrespective of the type of efficiency underlying the research question, an economic evaluation incorporates both costs (C) and health outcomes (O) (e.g., life-years gained; LYG) of the new technology compared to a standard technology defined by the current treatment guidelines [14]. The differences between treatment and intervention costs and health gains of the new intervention and the old intervention are calculated and its quotient is represented as the Incremental Cost-Effectiveness Ratio (ICER) [15], as shown in equation 1.1:

$$ICER = \frac{C_A - C_B}{O_A - O_B} \tag{1.1}$$

I. e., a new intervention might either be medically superior or inferior to the comparator and more costly or cost-saving resulting in four possible quadrants on the cost-effectiveness plane depicted in figure 1.1. While decisions about interventions in the quadrants II and IV are straightforward, decisions about interventions in the quadrants I and III need the definition of a maximal acceptable Willingness-To-Pay (WTP) threshold λ. Only if the ICER is below that threshold, the intervention should be reimbursed. Whether the value of the ICER directly decides if a new technology will be reimbursed depends on the country-specific administrative regulations, but it will at least have some influence on the decision of reimbursement. With $\lambda > 0$ the health care budget will be growing, but the growth can be governed by the absolute value of λ. A λ value of zero corresponds to a constant budget over time, i.e., new interventions will only be reimbursed if they are cost-saving. A negative λ will ensure a shrinkage of the health care budget.

It is also possible to calculate the Net Benefit (NB) of an intervention. This can take the form of the Net Monetary Benefit (NMB) by multiplying the value of the WTP threshold with the increment in the outcome $\Delta O = O_A - O_B$ and subtracting the costs differences $\Delta C = C_A - C_B$ [16]. For a WTP threshold of λ, the NMB is thus calculated as

$$NMB(\lambda) = \lambda \Delta O - \Delta C. \tag{1.2}$$

The Net Health Benefit (NHB) can be calculated accordingly by dividing the incremental costs ΔC by λ and substracting the result from the incremental effects ΔO, i.e.

$$NHB(\lambda) = \Delta O - \frac{\Delta C}{\lambda}. \tag{1.3}$$

If $NMB(\lambda) > 0$ or $NHB(\lambda) > 0$, the new intervention A should replace the old intervention B.

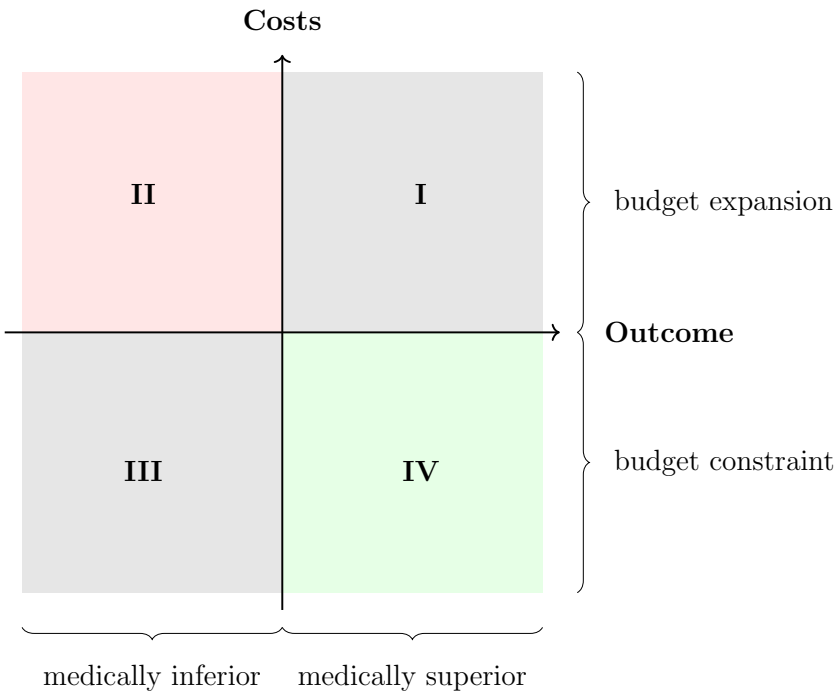

Figure 1.1: Cost-effectiveness plane. (Source: Own figure)

1.3 Decision analytic modeling techniques

Depending on the intervention at question, it might not be feasible or even possible to conduct studies observing all relevant data on costs and health outcomes and calculating the ICER or NMB in a timely manner for the decision. In this context modeling can be used to overcome the shortfalls of economic evaluations carried out alongside clinical trials. Modeling allows to include all available evidence on the effectiveness of a technology (not only from one clinical trial), a possible consideration of other comparators (i.e. the standard therapies) or settings not included in the clinical studies, and an appropriate time horizon [17]. HEDM uses a mathematical representation of a disease, in which parameters govern the disease activity or progression, the effectiveness of the interventions as well as costs and health outcomes linked to the disease states [18]. A wide range of modeling techniques are applied in the context of health economic evaluation with a more or less common taxonomy [19]. The following properties of models will be used for their categorization:

- **Entities:** The model population might either be modeled as a whole or by sub-groups of the total population of interest (*cohort model*) or might model each patient or person of the model population individually (*individual-based model*).
- **Time:** All models should have a defined time-period for which the model is built. But differences arise with regard to the handling of time, which can be *untimed*, *continuous* time or in *discrete* time steps. The time steps can be of *constant* or *variable* length.
- **Interaction:** In some models it is necessary to account for possible interactions of entities. Such *dynamic* models are either necessary in the modeling of infectious diseases, where the number of newly infected depends on the number of currently infected, or when entities compete for constraint resources (e.g., donor organs). In *static* models, entities have no influence on the course of the disease, the costs or the outcome of other entities.
- **Stochasticity:** Stochasticity regards the inclusion of randomness into a model. In *deterministic* models, parameter values are constant and the model results of two model runs will be identical. On the contrary, the values of model parameters in *stochastic* models are drawn from probability distributions and thus will return a different result for each model run. Therefore, stochastic models need to be run many times to calculate the expected (average) result.

Modeling Technique	Entities	Time	Interaction	Stochasticity
Decision Tree	Cohort	untimed	static	deterministic
Markov Model	Cohort	discrete, constant step length	static	deterministic
Differential Equations	Cohort	continuous	dynamic	deterministic
Individual Sampling Model	Individuals	discrete, constant step length	static	stochastic
Discrete Event Simulation	Individuals	discrete, variable step length	dynamic	stochastic
Agent-based modeling	Individuals	discrete, constant step length	dynamic	stochastic

Table 1.1: Overview of the different modeling techniques used in the context of health economic decision modeling.
(Source: Own table, based on Briggs et al. (2006) [18], Buxton et al. 1997 [17], Brennan et al. (2006) [19], Vynnycky et al. (2010) [20], Petrou et al. (2011) [21], Karnon et al. (2012) [22] and Gilbert et al. (2011) [23])

In principle every combination of the properties listed above is possible and even the same combination might have different names in the published literature. Additionally, mixtures of model types are applied, further contributing to the difficulty of an unambiguous distinction between them. However, not all models are used with equal frequency and the most common model types [21] are shown in Table 1.1 and will be described in more detail below.

Decision trees

The most simplistic model type used in health economics is the *decision tree*. Within decision trees the course of a disease is displayed as a hierarchical, one-directional tree starting with a decision node (■) and following with either a chance node (●) or a terminal node (◀) at the end of each branch (see Figure 1.2). A cohort of patients is followed through the model for each of the treatment choices, but with no interaction between the patients. Decision trees also only allow analyses for a fixed time frame, making it inapplicable for patients with diseases of varying length in a certain health state [19].

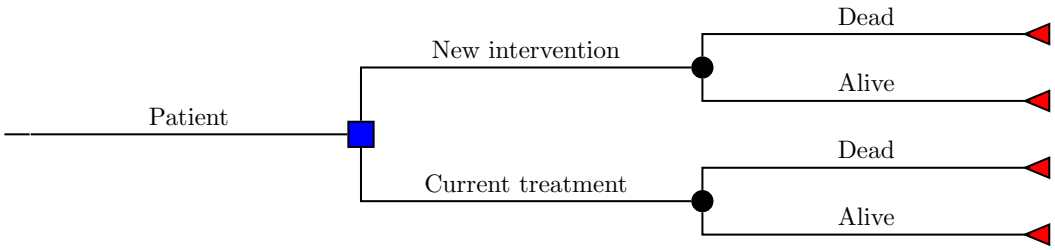

Figure 1.2: Basic structure of a decision tree model. (Source: Own figure)

Markov models

In health economic *Markov modeling* the course of a disease is described by discrete health states. Over time, patients of the model cohort move through or between these health states at the end of a cycle-length with a fixed time interval. The transition probabilities determine the amount of patients remaining in a health state and the amount of patients moving to other, connected health states at the end of a cycle. As the assumption that all patients progress exactly at the end of a cylce might be unjustified, a half cycle correction can be applied that shifts the time point of disease progression. However, there is no consensus on this method [24]. Markov models are usually run until all patients of a cohort reached an absorbing state (e.g. death) or

until a maximum number of cycles. The number of patients over all health states needs to stay constant in all cycles and patients are only allowed to be in one health state during a cycle [14]. Also there is no interaction modeled between the patients or parts of the cohorts. As the disease is modeled over health states, data need to be adjusted to accurately represent the clinical course, the costs and health outcomes at those fixed health states. Figure 1.3 shows two basic layouts of Markov models: Subfigure a) shows the model graph for a chronic disease, where a healthy patient who becomes sick cannot return to the healthy state. In the model layout for the acute disease (Subfigure b) a sick patient might become healthy again.

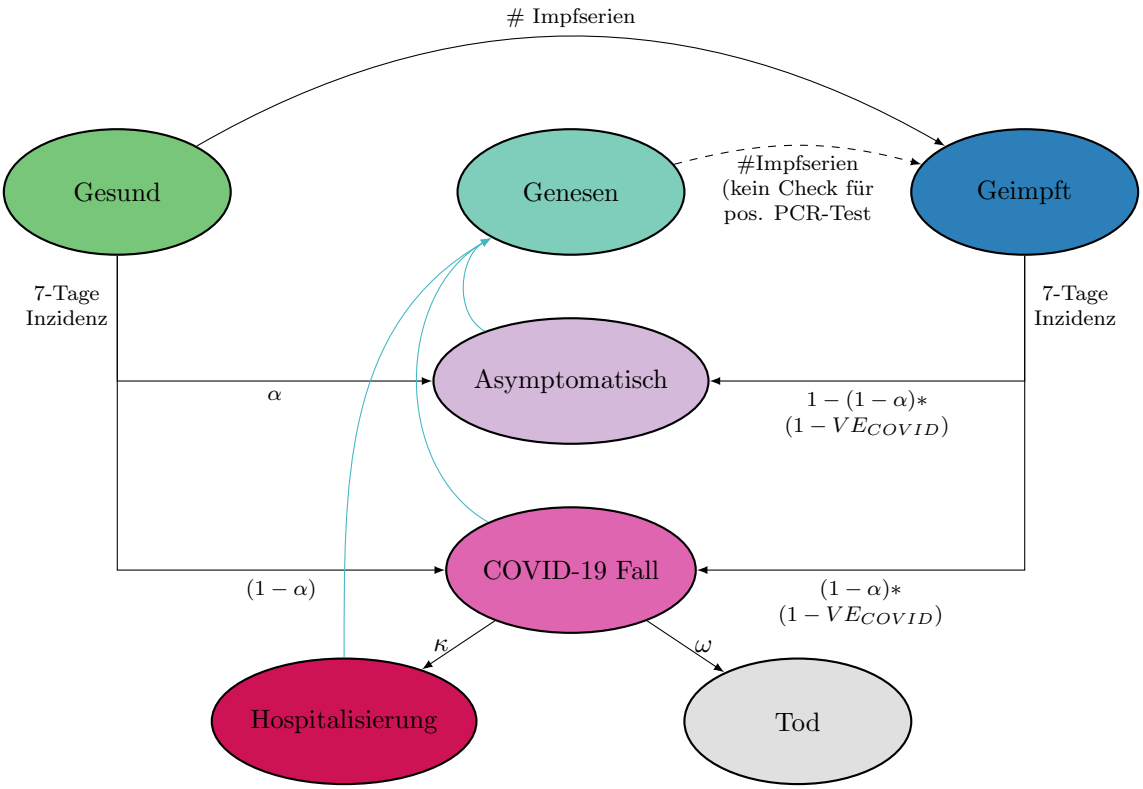

Figure 1.3: Basic structure of a markov model.
(Source: Own figure)

Dynamic transmission models

Markov models are able to model time more accurately than decision trees, bit it's inability to account for interactions between patients makes it inappropriate for infectious diseases. A

model type that still models cohorts but involves interactions between patients is the *dynamic transmission model*. This model type uses difference equations or Ordinary Differential Equations (ODEs) to calculate the number of persons in different health states – called compartments – over time. As these models are predominantly used to model infectious diseases, a common taxonomy has evolved that reflects the basic mechanic of the infectious disease. These basic compartments are "susceptible" (S), "pre-infectious"/"exposed" (E), "infectious" (I) or "recovered" (R) [20]. Depending on the natural history of the infection, model layouts can be described by combinations of those states, e.g., in a SIRS model healthy persons contract the infection and become immune for a period of time before they become healthy and susceptible for a new infection.

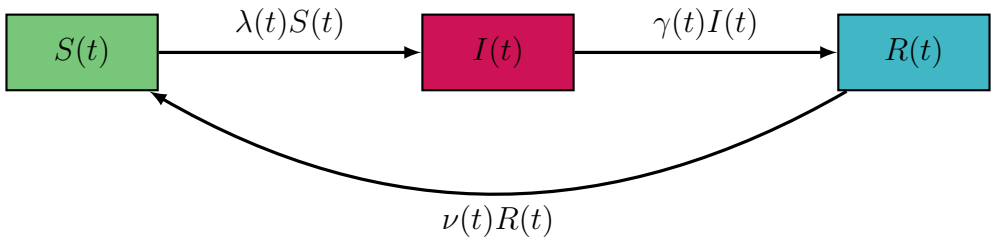

Figure 1.4: Basic structure of a dynamic transmission model based on ordinary differential equations. (Source: Own figure)

The number of new infections at time t is thereby dependent on the number of susceptible persons and the number of infected in the population at time t and the flow from one compartment to another is governed by corresponding rates as depicted in Figure 1.4. The parameters γ and ν represent the rate at which people become immune from the disease or become susceptible afterwards, respectively. The interaction between "infected" and "susceptible" is modeled by calculating the share of contacts of susceptible persons with infected persons. In combination with the probability of infection per contact, this term is called "the force of infection" (or λ). λ varies between infectious diseases and governs the number of susceptible persons becoming infected per unit of time. Costs and health outcomes of dynamic transmission models might be directly calculated for each health state or imputed by an additional decision tree or Markov-model. In the later cases it is important to reflect any changes in the health states – e. g. deaths – to the dynamic part of the overall model [20]. Advantages of dynamic transmission models are the ability to model long-term effects, effects resulting from interactions between entities (e. g. natural or vaccine-induced herd-immunity) and the effects of treatment of subgroups for the whole population. Disadvantages include the inability to capture the randomness of events and

allowing analyses for heterogeneity without a high effort [25].

Individual Sampling Models

Individual Sampling Models (ISMs) (or sometimes called *microsimulations* or *patient level simulations*) are basically decision trees or Markov models using single patients instead of cohorts [21]. Using a decision tree as an example, this means that for every chance node (●) the cohort is not split according to the probabilities of the branches following the chance node, but that for each patient it must be decided which branch he follows. Running the model for many patients will asymptotically result in the same model output as the cohort model, but ISM have the advantage that heterogeneity in patients can be integrated more easily. E.g., if the probability of a complication is linearly increasing with age, this can more easily and appropriately be reflected in an ISM compared to a cohort model which would need to include numerous numbers of subgroups.

Discrete Event Simulation

In contrast to the above described modeling techniques, Discrete Event Simulation (DES) focuses on the treatment pathway of a patient rather than modeling the course of the disease itself. Therefore entities (e. g., patients, health care professionals) with certain attributes (e. g., age, sex) experience certain events (e.g., disease progression, hospital admission) based on their attributes and, thereby, consume resources. An exemplified depiction of a DES can be found in Figure 1.5. In this example, a patient is diagnosed with a chronic disease for which a series of treatment options are available. The patient receives “Drug 1” and depending on the effectiveness stays in “State 1” of the disease for a certain period of time after which the treatment fails and the disease enters a phase of progression. After progression, the patient reveices “Drug 2”, which leads to a shorter period until the next progression, and so forth. The patient is followed until death and outcomes and costs are estimated along the way.

A strong emphasis of DES lies on the modeling of scarce resources and the behavior of entities competing or waiting for the availability of those resources. In this context interactions between entities can be modeled, as they can form waiting queues and access resources following predefined rules. Also, one advantage of DES is its capability of the inclusion of patient characteristics, thus accounting for heterogeneity [22]. Even though DES is one of the more advanced methods used in health economic evaluation it still has some disadvantages attached. Firstly, DES is rather process-oriented leaving out interactions between entities which are not

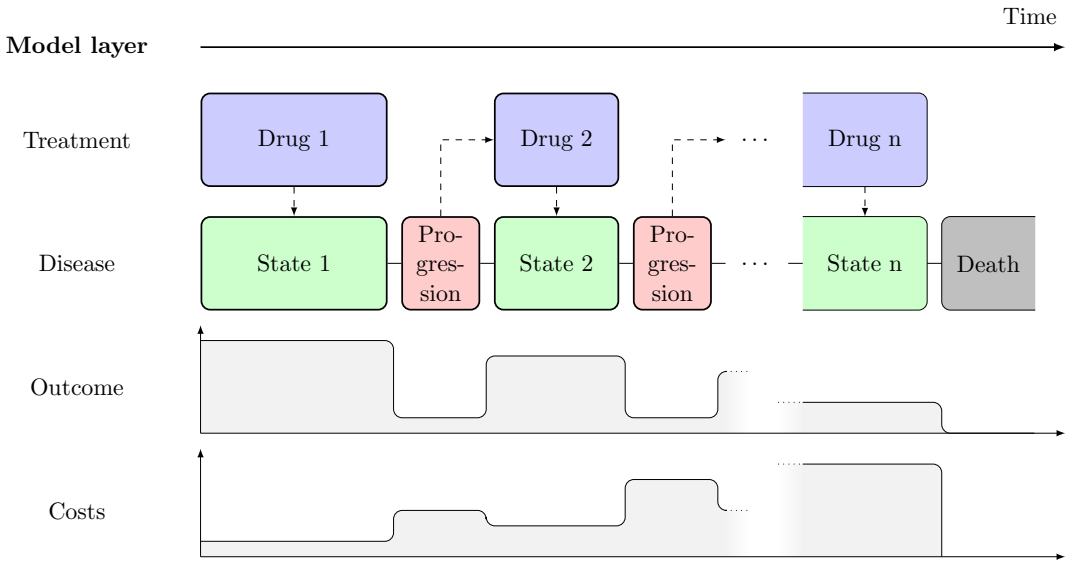

Figure 1.5: Basic structure of a discrete event simulation, showing a schematic patient flow without interaction. (Source: Own figure)

connected to resources, for example social contacts. Secondly, it leaves its entities in a rather passive role making them flow through the system instead of letting them decide independently [26].

Agent-based models

The last modeling technique presented here is the category of Agent-Based Models (ABMs). It has only recently become an option in health economic evaluation [27] while it is an already established modeling method in many other scientific fields since the 1990s [28]. ABM can be seen as an extension of DES, where the single persons are not modeled independently and run through the model subsequently, but all persons of a population are modeled at once, thus allowing for interactions between them [23]. Heterogeneity of patients can be included via different characteristics of an agent and conditioning the behavior during the model run on those characteristics. For example, in the modeling of Sexually Transmitted Diseases (STDs), sexual activity might depend on the age and the relationship status of the agents. This makes ABMs the modeling technique of choice if the population of interest consists of many different subgroups showing different types of behavior or if a disease is only persistent in a small group of individuals but occasionally spreads to the general population (so called “bridging”).

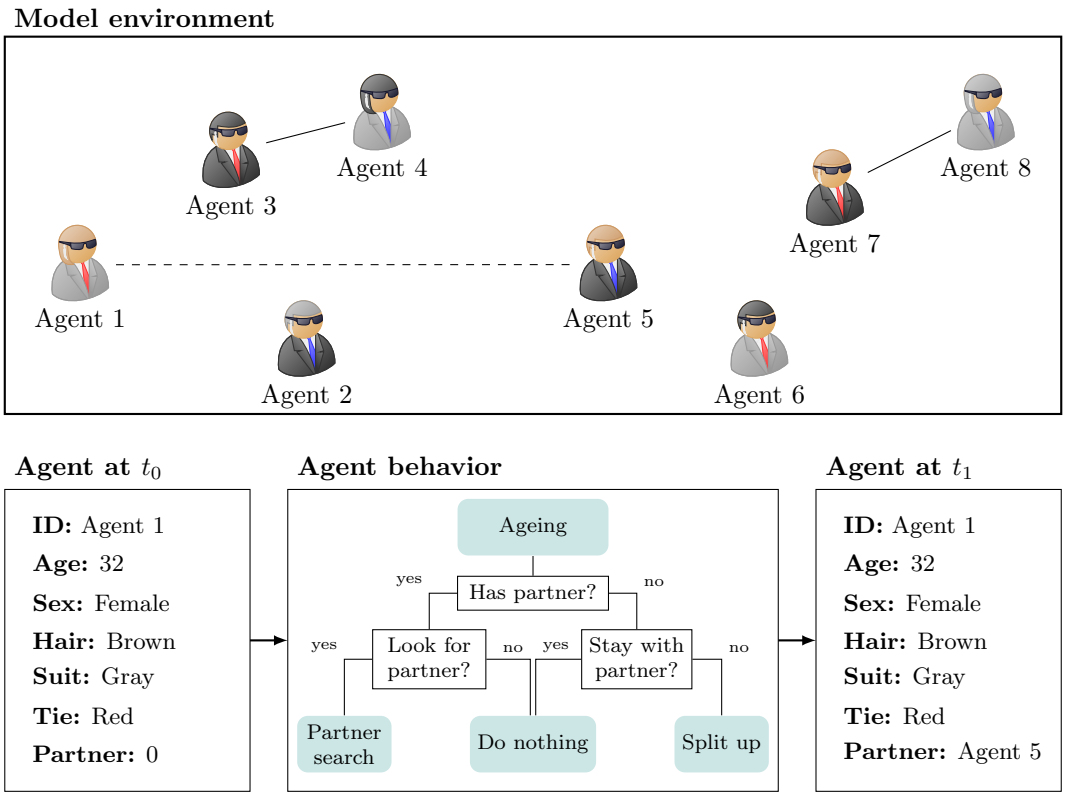

Figure 1.6: Basic structure of a agent-based model used for infectious disease modeling. (Source: Own figure)

Figure 1.6 shows a schematic depiction of an ABM. The hypothetical model consists of eight agents who are placed in a model environment with "Agent 3" and "Agent 4" as well as "Agent 7" and "Agent 8" being in a relationship at the simulation start (represented by the solid lines). Each agent has a set of characteristics, e.g., age, sex, the color of their hair, suit or tie. At each model iteration each agent runs through a set of functions – i.e., the agent behavior – shown in the bottom central of the figure. In this example, the agent ages and has a partner module where it either looks for a new partner, breaks up with it's current partner or stays in the current state (single or in partnership). The agent displayed in this example ("Agent 1" finds a new partner "Agent 5", dotted line) that matches her preferences for a new partner (different sex, same color hair, different color of tie and suit).

1.4 Uncertainty in decision analytic models

As all models are simplifications of reality, there are a few major uncertainties surrounding the most adequate way to depict a disease and possible interventions in models. The model results might be influenced by a number of choices during the process of modeling, ranging from the choice of the modeling technique itself, the parameter values entered into the model [18], as well as the layout and construction of the model. All three steps inherit the risk of omitting important factors biasing the results of the model and making them unsuitable for decision makers. For example, it may be inappropriate to model a disease with recurrent disease states using a decision tree, which only allows for one-directional progression through the model. Also, patient characteristics (e.g., age, socio-economic status) which strongly influence the effectiveness of interventions may not always be appropriately considered in certain types of models. At last, the true value of parameters might be unknown or parameter values are based on sample estimates. Giving these considerations, health economic models need to make allowance for these uncertainties. In the following an overview of the different categories of uncertainty will be given and the consequences of their dismissal on the model results and decision making will be discussed.

Type of uncertainty (other names)	Definition	Example	Regression equivalent
Heterogeneity	Differences between patients explainable by observable characteristics	Different effectiveness by age	β-coefficient
Variability (First-order, Monte Carlo error)	Random differences between patients identical in observed characteristics	Different effectiveness between identical patients	Error term ε
Statistical (Second-order)	Uncertainty in the estimation of parameter values from samples	Confidence Interval (CI) of effectiveness estimate	SE of β
Methodological	Differences in model results explainable by methodological choices	Value of discount rate	Type of regression
Structural (Model)	Differences in model results due to different model structures	Number of discrete disease states derived from CD4 cell count in Human Immunodeficiency Virus (HIV) patients	Number of explanatory variables X

Table 1.2: Overview of the different definitions of uncertainty used in the context of health economic decision modeling. (Source: Own table, based on Briggs et al. (2006) [18], Briggs et al. (2012) [29], Bojke et al. (2009) [30] and Bilcke et al. (2011) [31])

The formal categorization of sources of uncertainty differentiates into uncertainty around

values of input parameters, methodological uncertainty and structural uncertainty, although the terminology has been evolving over time [31, 29, 18]. A summary of the nomenclature found in the literature is given in Table 1.2. Parameter uncertainty surrounding modeling in HTA has been addressed for a relatively long time with the categorization of Briggs et al. (2006) being the most commonly used notation. The authors categorize parameter uncertainty in health economic models into *heterogeneity*, *variability* and *statistical uncertainty*. In this context *heterogeneity* is defined as the differences in model parameters (e.g., of treatment responses) between patients that can be explained by different, observable patient characteristics such as age, sex or weight. In contrast, *variability* is the influence of unobservable patient characteristics or unexplainable differences between otherwise identical persons. The term *statistical uncertainty* is used to describe the variance of estimated values of model parameters. Usually, mean values are used for model parameters, but this value might be uncertain as indicated by the CI or Standard Error (SE) of the estimated value.

This categorization has been amended by *methodological uncertainty* [31] and *structural uncertainty* [30, 29]. The first term refers to the choice of methodological aspects of cost-effectiveness analysis in general, e.g., the perspective of the analysis or the discount rate, and is not specific to modeling itself but applicable to all forms of economic evaluation. *Structural uncertainty* refers to the choice of the modeling technique as well as the structure or layout of the model. Different researchers might come to different conclusions which model type to use and how complex the model structure needs to be to include all relevant features of a disease or intervention.

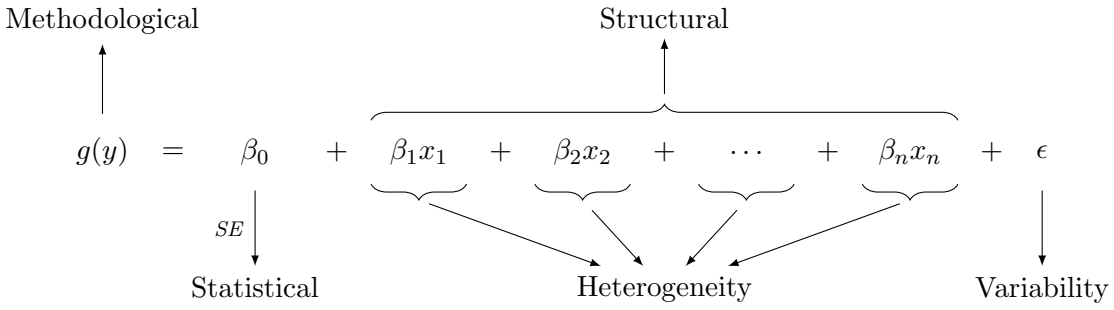

Figure 1.7: Regression analogy of the different types of uncertainty in health economic decision models. (Source: Own figure, based on Briggs et al. (2006) [18])

To illustrate the differences between the different types of uncertainties, Briggs et al. (2006) use regression analysis as an analogy. Figure 1.7 shows the different types of uncertainty as they would present themselves in a regression analysis. Heterogeneity would be the existence of

observable, explanatory variables X with corresponding β-coefficients. These variables explain differences in the outcome Y in addition to the intercept β_0. The remaining, unexplained variance of Y would be the variability, usually denoted as the error term ε. The statistical uncertainty would be the standard errors of all β coefficients or just the intercept β_0 if there is no heterogeneity. Structural uncertainty would be the number of explanatory variables with a detailed model including many, but possibly non-relevant, variables. Finally, methodological uncertainty would correspond to the specific regression used. Depending on the nature of the dependent variable Y, the type of regression analysis might range from Ordinary Least Squares (OLS) to more complicated Generalized Linear Models (GLMs) or Generalized Additive Models (GAMs).

The consequences of a failure to address uncertainties in HEDM may lead to the wrong conclusions by decision makers resulting in the non-optimal allocation of resources. Ignoring existing heterogeneity may limit the transferability of the results to other populations as the ICER may be dependent on patient characteristics as – for example – age or sex. It may also prohibit the most technically efficient use of resources [12], as limiting a new intervention to certain patient subgroups might prove cost-effective while the intervention would not be cost-effective for the whole population. Not addressing variability and statistical uncertainty may suggest a non-existing certainty of the result of the decision analysis to the decision makers. Failure to address structural uncertainty has several consequences. On the one hand, models can be too simple to capture all relevant aspects for the evaluation of the competing treatments. On the other hand, unnecessarily complex models are more difficult to interpret and increase proneness to coding mistakes and the amount data necessary to avoid assumptions. Lastly, not acknowledging methodological uncertainty leaves room for possible influence of the modeler on the model results and also limits the transferability to other settings.

1.5 Existing methods of dealing with uncertainty

There is a variety of methods and approaches available on how to deal with uncertainty in HEDM, and also how to present the results of decision models avoiding the possible consequences mentioned before. The first step to reduce uncertainties in modeling is the choice of the appropriate modeling technique. As different techniques might yield different outcomes, guidelines try to limit the influence of an arbitrary model choice on the results by outlining the prerequisites for each modeling technique [32]. Figure 1.8 gives an overview of the most important factors underlying the model choice.

The first important decision for the selection of a modeling technique is whether the model

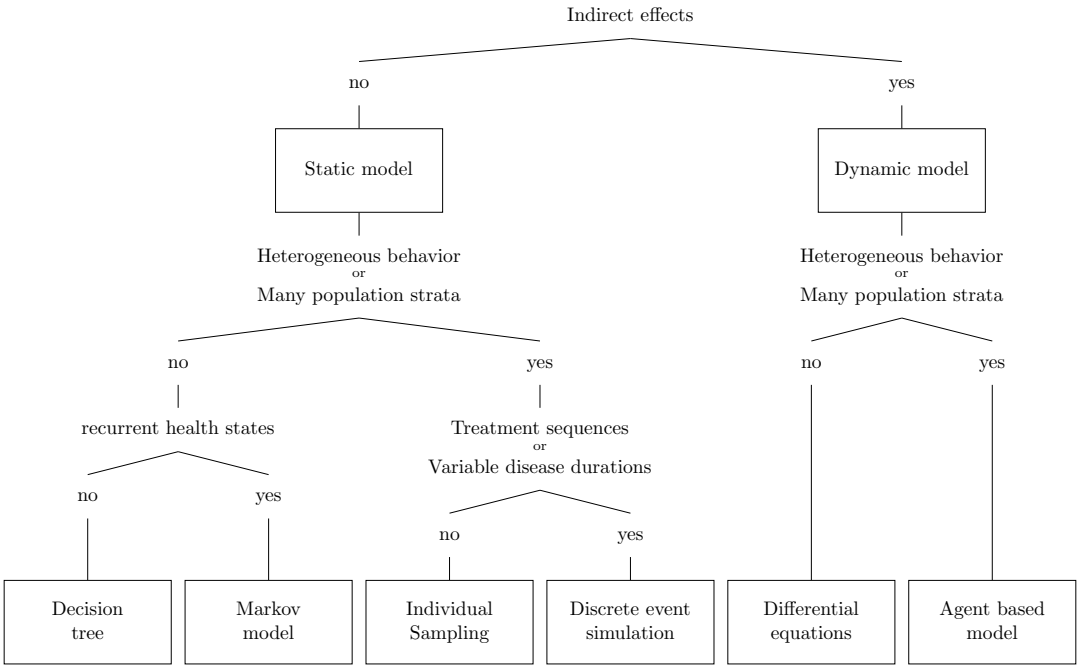

Figure 1.8: Flow chart depicting the choice of model type for a health economic decision model. (Source: Own figure, based on Roberts et al. (2012) [32])

should be capable of capturing indirect effects in the context of infectious diseases, i.e., if avoiding one case of a disease by the intervention leads to other cases avoided indirectly. If indirect effects can be expected, a dynamic model should be chosen. Otherwise a static model is sufficient. The second question is, if the model population shows heterogeneous behavior or if it consists of many strata, like age groups within different groups of sexual activity. For example, if disease progression or the intervention effectiveness are highly age-dependent, a cohort model with a population of average age might produce biased results. For dynamic models, the presence of heterogeneous behavior favors an ABM as this model can more efficiently incorporate many subgroups. If an intervention against an infectious disease in a homogeneous population is to be modeled, ODE-models should be used as they are less data-demanding as ABMs. In the field of static models the simplest model is a decision tree which is sufficient for a homogeneous population and a disease without any recurring health states. For chronic or recurrent diseases a Markov model is able to capture these disease features. If the course of the disease is significantly different between many patient groups or if the sequence of treatments or the infrastructure of health care provision plays a role in the evaluation of a new intervention, an ISM or a DES might be necessary [22]. There are some cases, in which no clear decision on the

model characteristics can be made – for example if there is no available evidence on a vaccine protecting against the carriage of a pathogen, indirect effects may not need to be incorporated [33]. Ideally, more than one model type is employed in this case, but if research funds are limited, the more complex model might be favorable as simpler models can be expressed as special cases of the more complex models.

Once a decision about the basic modeling technique has been made, the exact layout of the model needs to be determined. As stated above, *structural uncertainty* might arise from the decision of the exact layout of the model. One possible way to incorporate structural uncertainty is model averaging [34, 35]. This method allows to consider the results of different model layouts by calculating a weighted average of the single results of the respective layouts. In the method proposed by Jackson and colleagues (2009) [34], the weighting is achieved within a Bayesian framework by sampling from a posterior distribution of all models where the sample weight of each model is given by the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). This means, that the models with less parameters and better approximation of the observed data will contribute more to the averaged result. As Bojke et al. (2009) point out, the approach of model averaging is theoretically sound, but offers no advantages if there is no real-world data against which the model outputs can be compared [30].

With regard to *parameter* and *methodological uncertainty*, the main methodological approach are sensitivity analyses to incorporate uncertainty in the model [36]. The simplest form of sensitivity analysis is the Deterministic Sensitivity Analysis (DSA). Using this approach, values of single parameters, sets of parameters or methodological choices (e.g., discount rates) are varied within a certain range and the effect of this change on the ICER will be reported. For single parameters or sets of parameters (e.g., all cost data) the results are usually displayed as tornado diagrams where decision makers can easily identify and compare the impact of the changes of the parameter values on the ICER. The ranges of the parameter values are usually the CIs of the mean or minimum and maximum found in the sample from which the value was estimated. While DSA provides the respective ICERs for the lower and upper bound of parameters values, it provides no information on how likely those ICER values are. A further disadvantage is the limited number of combinations of parameters for which value ranges can be explored before the clarity of the presentation of the results is comprised.

Probabilistic Sensitivity Analysis (PSA) can overcome the weaknesses of DSA by drawing values for each parameter from an appropriate probability distribution for numerous times [18]. Depending on the distribution, values close to the mean value will be drawn more often and the extreme values may only occur occasionally. PSA also allows to explore the uncertainty around all model parameters at once leading to a empirical probability distribution of the ICER

itself. Another advantage is the inclusion of the covariance of parameters. For example, patients with high hospitalization costs may also be more likely to have higher costs for medication and may also show lower utility values. A disadvantage of PSA is the availability of several distributions with the same support (e.g., Poisson and negative binomial would be available for count date) and the necessity of detailed information on the distribution parameters (e.g., mean and variance). Additionally, the covariance of parameters may not always be readily available [29, 37]. Nevertheless, PSA has become a standard tool in HTA [38].

The presentation of the results of PSA is not as straightforward as for DSA as the recommended number of samples of the ICER is relatively high. Presenting all samples as a scatterplot on the cost-effectiveness plane as described in Figure 1.1 is providing only little information for decision makers. As PSA reflects a quasi-Bayesian method, the distribution of the ICER can be interpreted as a Cumulative Distribution Function (CDF) and thus for a given WTP threshold, the probability of $P(ICER < WTP)$ can be calculated corresponding to the probability of the new intervention being cost-effective. This probability can be plotted against a range of WTP-thresholds resulting in a Cost-Effectiveness Acceptability Curve (CEAC) [39]. In the presence of more than two alternatives, the probability of an intervention being most likely to be cost-effective among all interventions can be calculated for a given WTP threshold. Plotting only the intervention with the highest probability of all interventions against the thresholds of the WTP is called a Cost-Effectiveness Acceptability Frontier (CEAF). CEAC and CEAF allow decision makers to see how likely a certain intervention will be cost-effective depending on the WTP threshold and if the choice of the preferred – i.e., most likely to be cost-effective – intervention changes with a change in the WTP.

Calculating the net benefit for interventions for each draw in the PSA allows for another analysis useful for decision makers. For a given WTP threshold, the net benefit of the intervention with the highest probability can be calculated and subtracted from the net benefit of the maximal net benefit from all other available interventions. This difference is called Expected Value of Perfect Information (EVPI) and can be interpreted as the opportunity cost of further research to reduce uncertainty in the model [39]. Given the uncertainty around the model result, the choice of an intervention at a certain WTP threshold has a probability of having chosen the less cost-effective strategy as quantified by the CEAC. Decision makers could either go forward, choosing the intervention with the highest probability of being cost-effective, or they could postpone the decision and try to conduct further research to reduce the uncertainty and thus the risk of a wrong decision. From an economic perspective, decision makers should choose the option with the lower opportunity costs.

1.6 Aim and structure of the thesis

With increasing computational power and data availability, scientific models have become more complex and are thus capable to answer more detailed research questions. This is also true in HEDM where personalized healthcare (for example, with precision treatments in oncology [40]), complex interventions and behavioral aspects drive the complexity of health economic evaluations and models [41]. Complex models also call for methods to inform the increasing number of parameters with appropriate data. Otherwise, models become prone to be based on assumptions for the majority of their parameters, thus increasing the risk of bias. Methods from the emerging discipline of data science may offer possibilities to inform all parameters of complex health economics models [42].

The application of data science methods in health care can be divided into three main categories: Descriptive tasks, predictive tasks and causal inference (or counterfactual prediction) [43]. *Description* in a health care setting may for example be the estimation of incidence rates or simple frequencies of health-related behavioral variables. *Prediction* tries to link two variables observed in a population, e.g., calculating the risk of retinopathy for diabetic patients. Methods of *causal inference* try to estimate the causal relationship between variables, as it is for example necessary to estimate the effectiveness of interventions. Statistical methods are applied in all three tasks, e.g., by the calculation of confidence intervals for the corresponding estimators.

The aim of the current thesis is twofold: Firstly, statistical and data science methods will be explored that allow to incorporate the types of uncertainty described in section 1.4. Secondly, a new type of uncertainty will be described that arises from the increasing use of complex, individual-based stochastic model types – e.g., ABM and DES. As these models are not solved analytically but rather heuristically, the algorithms in those models may influence the model results. This is not covered by *methodological uncertainty*, which is dealing with the influence of normative methodological choices, nor by *structural uncertainty*, as the structure of the model (i.e., the layout of the health states) remains identical. Therefore, defining a new category of *algorithmic uncertainty* seems necessary for these complex models.

The cumulative part of the dissertation thesis consists of eight publications listed in Table 1.3. The first three publications deal with the classic parameter uncertainty and show how data for behavior in ABMs can be collected using web-mining, how claims data analysis can provide stratified cost data including the covariance of several cost categories and how heterogeneity can be accounted for in a decision model. The second chapter contains three studies that focus on methodological and structural uncertainty. The first article is a systematic review on the

Authors	Title	Published in
Scholz SM, Damm O, Elkenkamp S, Marcus U, Greiner W, Schmidt AJ	Population size and self-reported characteristics and sexual preferences of men-who-have-sex-with-men (MSM) in Germany based on social network data	PloS One 2019, Vol. 14(2):e0212175.
Scholz SM, Damm O, Schneider U, Ultsch B, Wichmann O, Greiner W	Epidemiology and cost of seasonal influenza in Germany - a claims data analysis	BMC Public Health 2019, Vol. 19:1090.
Scholz SM, Greiner W	An exclusive human milk diet for very low birth weight newborns – A cost-effectiveness and EVPI study for Germany	PloS One 2019, Vol. 14(12):e0226496.
Scholz SM, Mittendorf T	Modeling rheumatoid arthritis using different techniques - a review of model construction and results	Health Economic Review 2014, Vol. 4:18.
Treskova M, **Scholz SM**, Kuhlmann A	Cost Effectiveness of Elderly Pneumococcal Vaccination in Presence of Higher-Valent Pneumococcal Conjugate Childhood Vaccination: Systematic Literature Review with Focus on Methods and Assumptions	PharmacoEconomics 2019, Vol. 37(9):1093-1127.
Scholz SM, Koerber F, Meszaros K, Fassbender RM, Ultsch B, Welte RR, Greiner W	The cost-of-illness for invasive meningococcal disease caused by serogroup B Neisseria meningitidis (MenB) in Germany	Vaccine 2019, Vol. 37(12):1692-1701.
Geffen N, **Scholz SM**	Efficient and Effective Pair-Matching Algorithms for Microsimulations	The Journal of Artificial Societies and Social Simulation 2017, Vol. 20(3):8.
Geffen N[©], **Scholz SM**[©]	How various design decisions on matching individuals in relationships affect the outcomes of microsimulations of sexually transmitted infection epidemics	PloS One 2018, Vol. 13(8):e0202516.

[©] authors contributed equally

Table 1.3: Publications included in the cumulative dissertation thesis.

different modeling techniques used in modeling rheumatoid arthritis. The second publication is focusing on the effect of different structures and methodological choices for the evaluation of pneumococcal vaccines. The third publication shows the impact of methodological uncertainty in the calculation of costs for meningococcal disease. Chapter 4 introduces ”algorithmic uncertainty” as a source of uncertainty in complex ABMs. This is explored in two consecutive papers, with the first describing different matching algorithms for partner formation in STD modeling, followed by the second paper on the effect of these algorithms on the epidemiological model results. The results of all papers are summarized and discussed in Chapter 5. This includes overarching limitations across the publications and conclusions for the handling of uncertainties in modeling.

CHAPTER 2

Parameter Uncertainty: Heterogeneity, Variability and Stochasticity

2.1 Population size and self-reported characteristics and sexual preferences of men-who-have-sex-with-men (MSM) in Germany based on social network data

2.2 Epidemiology and cost of seasonal influenza in Germany – a claims data analysis

2.3 An exclusive human milk diet for very low birth weight newborns – A cost-effectiveness and EVPI study for Germany

CHAPTER 3

Methodological and structural uncertainty

3.1 The cost-of-illness for invasive meningococcal disease caused by serogroup B Neisseria meningitidis (MenB) in Germany

3.2 Modeling rheumatoid arthritis using different techniques – a review of model construction and results

3.3 Cost Effectiveness of Elderly Pneumococcal Vaccination in Presence of Higher-Valent Pneumococcal Conjugate Childhood Vaccination: Systematic Literature Review with Focus on Methods and Assumptions

CHAPTER 4

Algorithmic uncertainty

4.1 Efficient and Effective Pair-Matching Algorithms for Microsimulations

4.2 How various design decisions on matching individuals in relationships affect the outcomes of microsimulations of sexually transmitted infection epidemics

CHAPTER 5

Discussion

5.1 Summary of main results

The following section summarizes the main results of the eight publications of the cumulative dissertation. The publications will be put into the context of the respective state of the research and the contribution of the articles to the topic of uncertainty in health economic modeling will be highlighted.

Parameter Uncertainty

The first article reports on the population size, the self-reported characteristics and the sexual preferences of Men-who-have-Sex-with-Men (MSM) in Germany. For many reported Sexually Transmitted Infections (STIs) in Europe, MSM show an increasingly higher share of case numbers compared to Men-who-have-Sex-with-Women-only (MSW) since the beginning of the millennium [44]. However, knowing the size of the MSM population is crucial in order to correctly calculate standardized incidence or prevalence measures of diseases. In the context of MSM, this has been called the "denominator problem" [45]. Because of the increased risk of STIs, MSM are also of special interest to epidemiological and behavioral research and infectious disease modeling. Surveys focusing on MSM have been carried out repeatedly [46, 47], but stratifying sexual behavior data by age groups or regions within the group of MSM is difficult with the current sample sizes. Additionally, it is unclear to which extent survey participants reflect a self-selected group of persons with a high interest in sexual health, possibly leading to biased results.

The paper uses a novel approach to gather information from a website serving as an online dating portal and social network for MSM in Germany. Data from the website was publicly accessible and could be retrieved using the data science method of *web scraping* [48, 49]. Also called *data mining*, this method allows to retrieve seemingly unstructured information from a website and to transform it into machine-readable data-sets. The results of the study allow to incorporate a high level of *heterogeneity* in the behavioral aspects of infectious disease modeling, thus reducing the uncertainty of the effect of the behavior of different subgroups of the MSM population. The study results on the (marginal) distribution of MSM across federal state, age and sexual behavior provide crucial information for behavioral apects of STI modeling. The additional benefit of the analysis presented in this paper is the provision of the joint distribution of all these characteristics. E.g., the data-set allows to estimate the proportion of MSM between age 24 to 26 in Berlin who identify as bisexual, who prefer to be "more bottom" during intercourse and who are never performing safer sex. Having this highly stratified information available for modeling, allows to incorporate a very high degree of heterogeneity in a model and to find the optimal target group within the MSM community for interventions. This example also highlights the possible benefits of applying data science methods to make use of data from social networks and other internet sources, especially for populations which are otherwise hard to reach or where surveys are at risk of a sampling bias.

The second paper is also presenting an analysis in the field of infectious diseases and – in contrast to the behavioral aspects of the first paper – focuses on epidemiological aspects, resource use and costs. As already outlined in the section 1.3, the health economic evaluation of interventions against infectious diseases needs to incorporate indirect effects to fully account for the effectiveness of an intervention. In the evaluation of vaccines, finding the optimal target group for vaccination strategies becomes an increasingly important task. Vaccination campaigns are usually costly and for some diseases the subgroups of a population with the highest disease burden may not coincide with the subgroups with the highest incidence. This is the case for influenza, where children are affected by the highest incidence and hospital admission rates while the elderly and persons with comorbidities show the highest mortality [50]. The estimation of the cost-effectiveness of influenza immunization strategies therefore needs to be based on subgroup-specific data.

The analyses from this article take the above-mentioned issues into account and provide estimates for epidemiological and health economic model parameters from claims data of the largest German Statutory Health Insurance (SHI) covering the period from 2012 to 2014. The stratified analysis of the epidemiology and costs of influenza makes it possible to reduce uncertainty around costs and complication rates that is related to age. The study results cover the

full range of *heterogeneity* of the population with regard to age-specific complication rates and costs between different age-groups. The results from this paper have directly been used to estimate the cost-effectiveness of possible childhood vaccination strategies in Germany [51]. Due to the great sample size of the claims data-set, it was possible to incorporate detailed information for children and adolescents and to compare the cost-effectiveness of childhood influenza vaccination for multiple target groups, finding the technically most efficient vaccination strategy. The paper shows how epidemiological and health economic model parameters can be estimated for complex decision analytic models. The availability of the large claims data base allowed not only to address heterogeneity by stratifying the analyses by subgroups. It also allowed to quantify the variability by estimating not only the mean values but the complete probability distribution of the model parameters. Additionally, dependencies between model parameters could be quantified by the calculation of covariance matrices for the parameters values.

The last paper dealing with parameter uncertainty presents the results of a modeling study on the cost-effectiveness of an Exclusive Human Milk Diet (EHMD) for very low birth-weight newborns. Pre-term babies are at an increased risk of several complications including sepsis, Necrotizing Enterocolitis (NEC) or Broncho-Pulmonary Dysplasia (BPD). Recent studies have shown that substituting bovine milk-based products with human milk-based products in the feeding of very low birth-weight newborns has a protective effect against those complications [52, 53, 54, 55]. This may not only lead to a reduced life-long burden caused by the complications but may also reduce the treatment costs of pre-term babies.

The model consists of a decision tree where the treatment options are either a standard, bovine-based feeding or one of three different feeding strategies using an EHMD. The results of the cost-effectiveness analysis show the importance of the incorporation of *heterogeneity*: While the feeding strategy with EHMD for the overall population of very low birth-weight newborns is cost-effective, there is a strong difference in the ICER estimates for the different birth weight groups. The ICER is €15,293 for the lowest birth weight group of "<500gr" and increases with increasing birth weight up to the birth weight group "1250-1499gr" with a corresponding ICER of €72,402/Life-Year Gained (LYG). Using 1x the Gross Domestic Product (GDP) per capita as the WTP-threshold, including heterogeneity into the analysis results in EHMD to be a cost-effective intervention only for pre-term newborns with a birth-weight below 1000gr. The model also shows the usefulness of PSA to cover the other two types of parameter uncertainty, namely *variability* and *stochasticity*. Using appropriate probability distributions for the input parameters allows not only to calculate the probability of an ICER estimate but also supports decision makers in the allocation of research funds through the EVPI analysis.

Methodological and structural uncertainty

The first paper from chapter 3 is a model-based costing study for Invasive Meningococcal Disease (IMD) caused by serogroup B in Germany. IMD is a serious disease with a high mortality, leading to long-term physiological, neurological and psychological impairments among survivors [56, 57]. The incidence of IMD over all serogroups in Germany is 0.35 cases per 100,000 persons in 2017, thus qualifying for a rare disease [58]. The severity of the disease and it's low incidence pose challenges for the evaluation of vaccination programmes against IMD with regard to data collection: On the one hand high costs can be expected for a case, but the estimation of those costs using claims data is difficult as even large SHIs have only a few cases among their insured. On the other hand substantial parts of the resource consumption and the health effects may occur outside the period of data availability of claims data. One possible solution is the application of a model-based costing of a "typical" patient.

The paper follows this approach and uses public sources, expert opinion and published literature to estimate the cost input for the health economic evaluation of a vaccine against IMD caused by serogroup B. Direct and indirect costs are estimated along a hypothetical treatment pathway following 18 age groups of an average cohort of IMD patients between 2001 an 2016 in Germany. The article highlights the possible influence of *methodological uncertainty* on direct and indirect costs in health economic evaluation, especially in the case of long-term or lifelong resource use or health effects. As the results on costs of IMD show, especially indirect costs are strongly influenced by the different approaches for the calculation of the productivity losses. Changing the method for the calculation of indirect costs from the friction cost approach to the human capital approach leads to almost triple the costs at a three percent discount rate. Additionally changing the discount rate to zero percent leads to costs of €365,000 (using the human capital approach) representing an increase of factor 6.4. As these differences between the estimates show, different methodological choices can have a strong influence on the cost-effectiveness results. Scenario analyses – as presented in this paper – are therefore crucial to illustrate the consequences of those normative decisions to the decision makers.

The second paper of chapter 3 is a systematic literature review on the use of different modeling techniques in the field of Rheumatoid Arthritis (RA). The treatment of RA has been one of the most competitive areas for the development of new drugs, with treatment options including conventional Disease-Modifying Antirheumatic Drugs (DMARDs), biological and biosimiliar DMARDs, targeted synthetic DMARDs or glucocorticoid therapy [59]. As most treatments tend to loose their effectiveness in a patient over time, patients usually switch from one medication to the next. The high incidence of this chronic disease [60] in combination with the high

costs for the new medications lead to numerous cost-effectiveness studies. In this context, research questions of health economic models not only include the choice between the different treatment alternatives but also the possible treatment sequences of different medications.

The results presented in this article show the range of dealing with *structural uncertainty* in the literature, i.e., the variety of modeling techniques that are applied with a common aim in the context of cost-effectiveness modeling for treatment options against RA. While no direct conclusions can be made about the effect of the choice of the modeling technique on the ICER, the results of the review show the different depths of the analyses that are possible with the different model types: Decision trees are not capable to integrate the treatment sequences into the cost-effectiveness analysis. Markov models and ISM offer more possibilities of including questions around the timing of the different treatment options, but are to some extent limited by their fixed cycle length to account for the differences in the response times of the different medications. Only DES allows to fully incorporate different treatment sequences and patient characteristics into the cost-effectiveness analysis, thus allowing to answer detailed questions about the most cost-effective way of treating RA in different groups of patients.

The last paper of the chapter on methodological and structural uncertainty is a systematic review on the influence of methodological and structural assumptions on the cost-effectiveness results of health economic models for pneumococcal vaccination for the elderly. Pneumococcal disease can be caused by an infection with *Streptococcus pneumoniae* and leads to either Invasive Pneumococcal Disease (IPD) in the form of bacteremia or meningitis or to Non-Bacteremic Pneumococcal Pneumonia (NBPP). As with influenza, an important issue around the cost-effectiveness of pneumococcal vaccination strategies are indirect effects between children and the elderly. Routine childhood vaccination is in place in most developed countries mostly using Pneumococcal Conjugate Vaccines (PCVs) and is regarded an effective measure against pneumococcal infections [61]. Current health economic evaluations are looking at the cost-effectiveness of an additional vaccination programme for the elderly using either PCV or Pneumococcal Poly-Saccharide Vaccines (PPSVs), however with conflicting results.

The aim of the systematic review was to explain the differences in the results from the cost-effectiveness models with a focus on the methodological and structural choices made in those models. This included the assumptions made about the vaccine effectiveness – consisting of the absolute effectiveness estimate and the waning of the effectiveness over time – and the incidence of *S. pneumoniae* assumed for the evaluation of the elderly vaccination. This paper contributes to handling *structural uncertainty* by introducing a measure (i.e., the Expected Vaccine Protection Over Time (EVPOT)) that allows quantifying and comparing different structural choices regarding the vaccine effectiveness in models evaluating vaccine strategies. It also highlights

the importance of epidemiological assumptions, as unrealistic incidence rates in the base line scenario – especially for infectious diseases – might overestimate the potential impact of new interventions. Including these two aspects of *structural uncertainty* in the systematic review leads to a different conclusion with regard to the cost-effectiveness of an elderly pneumococcal vaccination programme than in a previous review, that did not include structural aspects and had found the elderly vaccination to be cost-effective [62].

Algorithmic uncertainty

The last two papers of the cumulative dissertation are part of the same project on the effect of the choice of the matching algorithm on the model results in ABMs modeling STIs. As already described in section 1.3, ABMs simulate single entities representing individual members of a society in contrast to the stratification of the population in compartments in ODE models. While in the latter, the number of contacts for disease transmission is calculated by simply multiplying each compartment with the contact matrix (i.e., the number of contacts of a compartment with every other compartment), ABMs need algorithms that specify explicitly which in-silico person had contact with which other in-silico persons at each model iteration. Ideally, the characteristics of agents (e.g., age, sex, sexual orientation) are incorporated in the algorithms and have an influence on who has contact with whom. As disease transmission is governed by these single contacts, the matching procedure is a core feature of the model with a significant impact on the spread of diseases. It is also important from a computational point of view. The matching procedure in an infectious disease ABM will be executed at each iteration and will usually involve the majority of agents, having a significant influence on the run-time of the simulation. Thus, having computationally efficient algorithms that effectively match agents according to their characteristics is a crucial part of an ABM.

The first of the two articles in this section describes the functioning of the six different algorithms. Two simulation experiments are conducted that compare the speed and the quality of the matches between agents. A distance function is defined to measure the quality of a possible match between two agents. For example, if persons would want a partner with the same age as themselves, the distance becomes greater as the age difference between the agents increases. An additional measure of effectiveness is the rank of the actual match of an agent among all possible agents. I.e., if an agent was matched with another agent but there would have been four agents (who were already matched with someone else) that would have had a lower distance, the rank for this agent is five. The mean and median rank over all agents was used as another quality measure. The aim of the second article was to analyze how the matching algorithms from the

first article translate into results of epidemiological infectious disease models. Additionally, the influence of multiple other decisions regarding the model structure on the model results were analyzed, including the population size, the influence of observable and unobservable characteristics on sexual behavior (i.e, heterogeneity and variability as described in section 1.4) and disabling the matching of agents who have formerly been in a relationship. To explore the influence of these design decisions, simulations were started with a fixed share of infected persons in the population and the spread of the infection was analyzed over time.

The results of this project contribute in two ways to the knowledge and methods of uncertainty in disease modeling. Firstly, the results provide information on the influence of behavioral *heterogeneity* and *variability* on the epidemiological model output of ABMs. As outlined in section 1.3, ABMs allow a more detailed, explicit modeling of behavior driving the spread of infectious diseases. In this context, it is important to analyze the necessary level of detail in the agents' behavior and the influence on the model results. Secondly, the studies provide insides on a new type of uncertainty arising from the use of the novel approach of ABMs in infectious disease modeling. As it is no longer possible to incorporate contacts between individuals in an abstract, implicit way as it is done in ODE models, more explicit ways of matching agents need to be employed. This can be done in several ways, creating uncertainty around the choice of the matching algorithm. Researchers should be aware of this *algorithmic uncertainty* and reduce the uncertainty of their choice of algorithm by sensitivity analyses using different algorithms.

5.2 Limitations

The results included in this dissertation need to be seen in light of several restrictions. As the limitations of each single study are discussed at length within the articles, the following section presents overall issues concerning the generalization and the transferability of the results.

One overarching limitation of the thesis as a whole is that the different aspects of uncertainty have not all been addressed in a single project or rather one model. This makes it difficult to quantify the absolute or relative impact that different types of uncertainty have on the model results. However, as there are many model types and countless fields of applications, it is not possible to cover all aspects of uncertainty within a single model. For example, uncertainties surrounding behavioral aspects play a major role in infectious disease modeling. On the contrary, behavioral aspects are widely disregarded in cost-effectiveness analyses of new therapeutics for non-communicable diseases. In most cases, it is only indirectly included via the efficacy of new, behavioral interventions like smoking cessation or weight-loss, but not modeled explicitly.

Furthermore, with eight articles included, the thesis covers a broad range of diseases and model types.

The data science methods presented in the first two articles of the first chapter are able to provide data for a high level of *heterogeneity* in decision models. However, the access to databases – e.g., online social networks or claims data – is not always available or may be restricted. It is not allowed to scrape data from websites if the terms and conditions of the website prohibit the use of scripts and if the data of interest is only accessible for registered and logged-in users. Furthermore, there are unresolved legal issues around the usage of web-scraping to download structured or unstructured information from websites [63, 64]. However, more and more websites are providing Application Program Interfaces (APIs) giving users and scientists controlled access to their data (e.g., Twitter provides an API [65]). Big social networks also increasingly employ their own research teams allowing other researchers to cooperate and to use their data. Researchers nevertheless need to be careful of self-selection biases in social networks. These biases might limit the representativeness of the extracted data [66].

A limitation with regard to claims data of SHI in Germany is, that data from social security institutions are under strict data-protection regulations [67]. But SHI data are in principle available for research purposes if a researcher finds a SHI which is willing to provide the data [68]. As the client structure of each of the SHIs is historically grown there might be a systematic difference between the members of the different SHIs leading to a potential bias when using only the data from one SHI. This might be avoidable in the future, as efforts have been undertaken from a political side to increase the accessibility of claims data to researchers by creating a central institution providing access to the data of all SHIs [69]. Furthermore, claims data come with a series of shortcomings as they are not primarily collected for research purposes. Among those disadvantages are the unspecific documentation of outpatient diagnoses [70], the missing clinical information, no information on services which are not reimbursable [71] as well as up-coding and incomplete documentation. Another limitation is the difference of the insured between the different SHI and the difference between the statutorily insured and the privately insured population [72, 73].

5.3 Conclusions

The topic of the dissertation is the identification and the handling of uncertainties in decision analytic modeling. In chapter 1 the necessity of HTA and the role of modeling within HTA were established and categories of uncertainties surrounding HEDM have been defined. Also, exist-

ing methods have been described that allow the quantification and illustration of uncertainties in health economic models to decision makers. Chapter 2 presented three articles with the subject of the three parameter uncertainties *heterogeneity*, *variability* and *statistical uncertainty*. The employed data science methods present a novel approach for generating input parameters for behavioral aspects as well as for epidemiological and health economic parts of decision models. Especially web-scraping can be seen as an useful, novel tool to generate data in the area of health economic evaluation. The impact of methodological and structural uncertainty was shown in Chapter 3. The presented results made it apparent how important sensitivity analyses are to quantify the impact of normative decisions on the study results. The publications also showed the relevance of structural uncertainty with regard to the range of modeling techniques applied to the same research question and the influence of the choices of the model layout using the same modeling technique. In the last Chapter 4, two articles were presented highlighting the possible impact of a new type of uncertainty – the *algorithmic uncertainty* – emerging in complex, individual-based models.

As models are moving from cohort models towards individual-based models, model results cannot be calculated directly and analytical solutions are replaced by heuristic algorithms. For many of these heuristic tasks, a variety of algorithms is available [74, 75] (for example, there are 43 different algorithms available for the simple task of "sorting" an array). This opens up a new category of uncertainty as the model structure remains the same and methodological uncertainty covers only normative decisions about the methods of economic evaluations. The results of this thesis offer a first insight on algorithmic uncertainty and shows the influence of the choice of the matching algorithm in the field of infectious disease modeling. Further research is necessary to explore the extent of algorithmic uncertainty in other models and disease areas.

Health care systems in developed and developing countries are confronted with more complex questions on how to distribute resources and provide tailored health care services in an efficient way. HEDM needs to reflect the complexity of those questions if it wants to remain a useful tool to guide decision makers on these issues. For example, vaccine-hesitant or anti-vaccine behavior is listed as one of the ten greatest threads to human health by WHO [76]. Models on the cost-effectiveness of vaccines need to incorporate this phenomenon to correctly model outbreaks of non-endemic diseases like measles. In contrast to "classical" observable characteristics like age or sex, vaccine hesitancy is an unobservable characteristic. Therefore new models need to be developed, which are able to incorporate such personal preferences. Modeling HIV and other STIs in the MSM community is another example for the need of more complex models. Two factors have changed the sexual behavior of parts of the MSM community: The availability of Pre-Exposure Prophylaxis (PrEP) reduces the risk of infection with HIV during unprotected sex

and online dating allows persons to find sexual partners with the same HIV-status. Both factors make condom use for the protection against HIV obsolete, but other STDs – e.g., syphilis – show an increase in their incidence [77]. Modeling these interactions between the treatment of one disease (i.e., HIV) and the incidence of other diseases (i.e., syphilis) in a specific community (i.e., MSM) cannot be done with "classical" Markov models or decision trees. Rather, applying more complex ABMs and using data sources like online social networks allows to adequately model the cost-effectiveness of PrEP. Scientists need to analyze the uncertainties arising from new modeling techniques as well as the limitations of new data sources and need to properly communicate these shortcomings to decision makers. This dissertation provides a first step into this direction and shows how to deal with uncertainty in modeling using statistical and data science methods.

BIBLIOGRAPHY

[1] World Health Organization. Basic documents. 48th ed. World Health Organization; 2014.

[2] Baltagi BH, Lagravinese R, Moscone F, Tosetti E. Health Care Expenditure and Income: A Global Perspective. Health Economics. 2017;26(7):863-74.

[3] Wang F. More Health Expenditure, Better Economic Performance? Empirical Evidence From OECD Countries. INQUIRY: A Journal of Medical Care Organization, Provision and Financing. 2015;52.

[4] Hall R, Jones C. National Bureau of Economic Research, editor. The Value of Life and the Rise in Health Spending. Cambridge, MA: National Bureau of Economic Research; 2004.

[5] Darvas Z, Moes N, Myachenkova Y, Pichler D. bruegel, editor. The macroeconomic implications of healthcare. bruegel; 2018.

[6] Organization for Economic Cooperation and Development. Value for money in health spending. OECD health policy studies. Paris: OECD; 2010. Available from: `http://paperc.de/10899-value-for-money-in-health-spending-9789264088818`.

[7] Savedoff WD. What should a country spend on health care? Health affairs (Project Hope). 2007;26(4):962-70.

[8] Jenkner, Eva, Leive, Adam. International Monetary Fund, editor. Health Care Spending Issues In Adanced Economies. International Monetary Fund; 2010.

[9] Banta D, Jonsson E. History of HTA: Introduction. International Journal of Technology Assessment in Health Care. 2009;25 Suppl 1:1-6.

[10] Perleth M, Gibis B, Göhlen B. A short history of health technology assessment in Germany. International Journal of Technology Assessment in Health Care. 2009;25 Suppl 1:112-9.

[11] Perleth M, Busse R, editors. Health technology assessment: Konzepte, Methoden, Praxis für Wissenschaft und Entscheidungsfindung. 2nd ed. Berlin: MWV Med. Wiss. Verl.-Ges; 2014.

[12] Palmer S, Torgerson DJ. Economic notes: definitions of efficiency. BMJ (Clinical research ed). 1999;318(7191):1136.

[13] Bhatt J, Jahnke N, Smyth AR. Once-daily versus multiple-daily dosing with intravenous aminoglycosides for cystic fibrosis. The Cochrane Database of Systematic Reviews. 2019;9:CD002009.

[14] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford University Press; 2005.

[15] Schöffski O, von der Schulenburg JM, editors. Gesundheitsökonomische Evaluationen. Dordrecht: Springer; 2012.

[16] Zethraeus N, Johannesson M, Jönsson B, Löthgren M, Tambour M. Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. PharmacoEconomics. 2003;21(1):39-48.

[17] Buxton MJ, Drummond MF, van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in Ecomomic Evaluation: An Unavoidable Fact of Life. Health Economics. 1997;6(3):217-27.

[18] Briggs AH, Claxton K, Sculpher MJ. Decision modelling for health economic evaluation. Handbooks in health economic evaluation series. Oxford: Oxford University Press; 2011.

[19] Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Economics. 2006;15(12):1295-310.

[20] Vynnycky E, White RG. An introduction to infectious disease modelling. Oxford University Press; 2010.

[21] Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. BMJ (Clinical research ed). 2011;342:d1766.

[22] Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–4. Value in Health. 2012;15(6):821-7.

[23] Gilbert GN, Troitzsch KG. Simulation for the social scientist. 2nd ed. Maidenhead: Open Univ. Press; 2011.

[24] Naimark DMJ, Kabboul NN, Krahn MD. The half-cycle correction revisited: redemption of a kludge. Medical Decision Making. 2013;33(7):961-70.

[25] Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes : a focused review of modelling approaches. PharmacoEconomics. 2008;26(3):191-215.

[26] Siebers PO, Macal CM, Garnett J, Buxton D, Pidd M. Discrete-event simulation is dead, long live agent-based simulation! Journal of Simulation. 2010;4(3):204-10.

[27] Chhatwal J, He T. Economic evaluations with agent-based modelling: an introduction. PharmacoEconomics. 2015;33(5):423-33.

[28] Heath B, Hill R, Ciarallo F. A Survey of Agent-Based Modeling Practices (January 1998 to July 2008). Journal of Artificial Societies and Social Simulation. 2009;12(4).

[29] Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–6. Value in Health. 2012;15(6):835-42.

[30] Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing structural uncertainty in decision analytic models: a review and application of methods. Value in Health. 2009;12(5):739-49.

[31] Bilcke J, Beutels P, Brisson M, Jit M. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. Medical Decision Making. 2011;31(4):675-92.

[32] Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Medical Decision Making. 2012;32(5):678-89.

[33] Christensen H, Irving T, Koch J, Trotter CL, Ultsch B, Weidemann F, et al. Epidemiological impact and cost-effectiveness of universal vaccination with Bexsero(®) to reduce meningococcal group B disease in Germany. Vaccine. 2016;34(29):3412-9.

[34] Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. Journal of the Royal Statistical Society Series A, (Statistics in Society). 2009;172(2):383-404.

[35] Jackson CH, Bojke L, Thompson SG, Claxton K, Sharples LD. A framework for addressing structural uncertainty in decision models. Medical Decision Making. 2011;31(4):662-74.

[36] Saltelli A. Sensitivity analysis in practice: A guide to assessing scientific models. Reprinted. ed. Hoboken, NJ: Wiley; 2005.

[37] Briggs A, Nixon R, Dixon S, Thompson S. Parametric modelling of cost data: some simulation evidence. Health Economics. 2005;14(4):421-8.

[38] Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Economics. 2005;14(4):339-47.

[39] Barton GR, Briggs AH, Fenwick EAL. Optimal cost-effectiveness decisions: the role of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value in Health. 2008;11(5):886-97.

[40] Garraway LA, Verweij J, Ballman KV. Precision Oncology: An Overview. Journal of Clinical Oncology. 2013;31(15):1803-5.

[41] Dzau VJ, Ginsburg GS. Realizing the Full Potential of Precision Medicine in Health and Health Care. JAMA. 2016;316(16):1659-60.

[42] Blei DM, Smyth P. Science and data science. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(33):8689-92.

[43] Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. CHANCE. 2019;32(1):42-9.

[44] Spiteri G. Sexually transmitted infections in Europe 2013. Sexually transmitted infections in Europe 2013. [Luxembourg]: [Publications Office]; 2015.

[45] Marcus U, Schmidt AJ, Hamouda O, Bochow M. Estimating the regional distribution of men who have sex with men (MSM) based on Internet surveys. BMC Public Health. 2009;9:180.

[46] Drewes J, Kruspe M. Schwule Männer und HIV/AIDS 2013: Schutzverhalten und Risikomanagement in den Zeiten der Behandelbarkeit von HIV. vol. 61 of Aids-Forum DAH. Berlin: Deutsche AIDS-Hilfe e.V; 2016.

[47] The EMIS Network. EMIS 2010: the European Men-Who-Have-Sex-With-Men internet survey: Findings from 38 countries. Technical report. Stockholm: European Centre for Disease Prevention and Control; 2013.

[48] Broucke Sv, Baesens B. Practical Web Scraping for Data Science: Best Practices and Examples with Python. Berkeley, CA: Apress; 2018.

[49] Mitchell R. Web scraping with Python: Collecting more data from the modern web. 2nd ed. Beijing and Boston and Farnham and Tokyo and Sebastopol: O'Reilly; 2018.

[50] Cromer D, van Hoek AJ, Jit M, Edmunds WJ, Fleming D, Miller E. The burden of influenza in England by age and clinical risk group: a statistical analysis to inform vaccine policy. The Journal of Infection. 2014;68(4):363-71.

[51] Scholz SM. Cost-effectiveness of routine childhood vaccination against seasonal influenza in Germany. Value in Health. 2020;[submitted].

[52] Abrams SA, Schanler RJ, Lee ML, Rechtman DJ. Greater mortality and morbidity in extremely preterm infants fed a diet containing cow milk protein products. Breastfeeding Medicine. 2014;9(6):281-5.

[53] Arslanoglu S, Corpeleijn W, Moro G, Braegger C, Campoy C, Colomb V, et al. Donor human milk for preterm infants: current evidence and research directions. Journal of Pediatric Gastroenterology and Nutrition. 2013;57(4):535-42.

[54] Hair AB, Bergner EM, Lee ML, Moreira AG, Hawthorne KM, Rechtman DJ, et al. Premature Infants 750-1,250 g Birth Weight Supplemented with a Novel Human Milk-Derived Cream Are Discharged Sooner. Breastfeeding Medicine. 2016;11:133-7.

[55] O'Connor DL, Kiss A, Tomlinson C, Bando N, Bayliss A, Campbell DM, et al. Nutrient enrichment of human milk with human and bovine milk-based fortifiers for infants born weighing <1250 g: a randomized clinical trial. The American Journal of Clinical Nutrition. 2018;108(1):108-16.

[56] Edmond K, Clark A, Korczak VS, Sanderson C, Griffiths UK, Rudan I. Global and regional risk of disabling sequelae from bacterial meningitis: a systematic review and meta-analysis. The Lancet Infectious Diseases. 2010;10(5):317-28.

[57] Robert Koch-Institut. Invasive Meningokokken-Erkrankungen 2012-2015. Epidemiologisches Bulletin. 2016;(43).

[58] Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, et al. Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. Value in Health. 2015;18(6):906-14.

[59] Kerschbaumer A, Sepriano A, Smolen JS, van der Heijde D, Dougados M, van Vollenhoven R, et al. Efficacy of pharmacological treatment in rheumatoid arthritis: a systematic literature research informing the 2019 update of the EULAR recommendations for management of rheumatoid arthritis. Annals of the Rheumatic Diseases. 2020.

[60] Minichiello E, Semerano L, Boissier MC. Time trends in the incidence, prevalence, and severity of rheumatoid arthritis: A systematic literature review. Joint Bone Spine. 2016;83(6):625-30.

[61] Tin Tin Htar M, Christopoulou D, Schmitt HJ. Pneumococcal serotype evolution in Western Europe. BMC Infectious Diseases. 2015;15:419.

[62] Porchia BR, Bonanni P, Bechini A, Bonaccorsi G, Boccalini S. Evaluating the costs and benefits of pneumococcal vaccination in adults. Expert Review of Vaccines. 2017;16(2):93-107.

[63] Krotov V, Silva L, editors. Legality and Ethics of Web Scraping; 2018.

[64] Bernard B. Web Scraping and Crawling Are Perfectly Legal, Right?; 2018. Available from: https://benbernardblog.com/web-scraping-and-crawling-are-perfectly-legal-right/.

[65] Twitter I. API reference index; 2020. Available from: https://developer.twitter.com/en/docs/api-reference-index.

[66] Mellon J, Prosser C. Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. Research & Politics. 2017;4(3):205316801772000.

[67] Deutscher Bundestag. Sozialgesetzbuch Erstes Buch; 11. Dezember 1975.

[68] Deutscher Bundestag. Sozialgesetzbuch Fünftes Buch; 19. November 2019.

[69] Deutscher Bundestag. Gesetzes für eine bessere Versorgung durch Digitalisierung und Innovation: Digitale-Versorgung-Gesetz – DVG; 18. Dezember 2019.

[70] Hoffmann F, Andersohn F, Giersiepen K, Scharnetzky E, Garbe E. Validierung von Sekundärdaten. Grenzen und Möglichkeiten. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz. 2008;51(10):1118-26.

[71] Zeidler J. Potentiale und Limitationen von GKV-Routinedaten. Hannover : Gottfried Wilhelm Leibniz Universität Hannover;.

[72] Hoffmann F, Icks A. Unterschiede in der Versichertenstruktur von Krankenkassen und deren Auswirkungen für die Versorgungsforschung: Ergebnisse des Bertelsmann-Gesundheitsmonitors. Das Gesundheitswesen. 2012;74(5):291-7.

[73] Jaunzeme J, Eberhard S, Geyer S. Wie "repräsentativ" sind GKV-Daten? Demografische und soziale Unterschiede und Ähnlichkeiten zwischen einer GKV-Versichertenpopulation, der Bevölkerung Niedersachsens sowie der Bundesrepublik am Beispiel der AOK Niedersachsen. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz. 2013;56(3):447-54.

[74] Knuth DE. The Art of Computer Programming, Volume 2: Seminumerical Algorithms. 2nd ed. Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc; 1997.

[75] Knuth DE. The Art of Computer Programming, Volume 3: Sorting and Searching. 2nd ed. Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc; 1998.

[76] World Health Organization. Report of the SAGE Working Group on Vaccine Hesitancy;. Available from: https://www.who.int/immunization/sage/meetings/2014/october/SAGE_working_group_revised_report_vaccine_hesitancy.pdf?ua=1.

[77] Jansen K, Schmidt AJ, Drewes J, Bremer V, Marcus U. Increased incidence of syphilis in men who have sex with men and risk management strategies, Germany, 2015. Euro Surveillance. 2016;21(43).

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Doktorarbeit selbständig verfasst und gelieferte Datensätze, Zeichnungen, Skizzen und graphische Darstellungen selbständig erstellt habe. Ich habe keine anderen Quellen als die angegebenen benutzt und habe die Stellen der Arbeit, die anderen Werken entnommen sind – einschließlich verwendeter Tabellen und Abbildungen – in jedem einzelnen Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht.

_____ _____

Ort, Datum Stefan Scholz

Erklärung über frühere Promotionsversuche

Ich erkläre hiermit, dass ich bisher noch keinen Promotionsversuch an einer in- und ausländischen Hochschule unternommen habe und ich an keiner anderen Fakultät oder Hochschule als Doktorand angenommen bzw. zur Promotion zugelassen worden bin.

Ort, Datum

Stefan Scholz

Gedruckt auf alterungsbeständigem Paper °° ISO 9706