

## Working Paper Modelling

Version 1.0 – 19/11/2009

**Contact:**

Institute for Quality and Efficiency in Health Care (IQWiG)  
Dillenburger Straße 27  
D-51105 Köln

Tel.: +49-221/35685-0

Fax: +49-221/35685-1

E-Mail: [knb-methoden@iqwig.de](mailto:knb-methoden@iqwig.de)

## Table of contents

List of abbreviations.....	iii
<b>1 Background.....</b>	<b>1</b>
<b>2 Definition of the term “model” .....</b>	<b>2</b>
<b>3 Aim of health economic modelling.....</b>	<b>3</b>
<b>4 Process for developing a model .....</b>	<b>4</b>
<b>5 Influence diagrams .....</b>	<b>6</b>
5.1 Definition and purpose .....	6
5.2 Components.....	6
<b>6 Model concept.....</b>	<b>7</b>
6.1 Definition and purpose .....	7
6.2 Components.....	7
6.3 Validation .....	8
<b>7 Features of the model.....</b>	<b>9</b>
7.1 Appropriate depth.....	9
7.2 Perspective .....	9
7.3 Time horizon .....	10
7.4 Discounting .....	10
<b>8 Assessment of data .....</b>	<b>11</b>
8.1 Objectives.....	11
8.2 Potential sources.....	12
8.2.1 Clinical trials .....	12
8.2.2 Epidemiological studies .....	13
8.2.3 Automatically generated databases .....	13
8.2.4 Registries.....	13
8.2.5 Compiled statistics .....	14
8.2.6 Expert opinion.....	14
<b>9 Functional relations.....</b>	<b>15</b>
9.1 Definition .....	15
9.2 Purpose.....	15
9.3 General concepts .....	15
<b>10 Modelling techniques .....</b>	<b>18</b>
10.1 Choice of modelling technique .....	18

---

10.2	Aspects in the choice of modelling technique.....	18
10.2.1	Cohort versus individual models.....	19
10.2.2	Event-based versus state transition-based simulation .....	19
10.3	Techniques .....	20
10.3.1	Decision trees .....	20
10.3.2	Discrete event simulation.....	20
10.3.3	Markov models.....	23
10.3.4	Agent-based simulation.....	26
10.3.5	Transmission models.....	26
<b>11</b>	<b>Handling variability and uncertainty .....</b>	<b>28</b>
11.1	Types of variability .....	28
11.2	Quantifying uncertainty.....	28
11.3	Handling uncertainty and its effects on model results .....	29
<b>12</b>	<b>Validation.....</b>	<b>31</b>
<b>13</b>	<b>Model documentation .....</b>	<b>33</b>
<b>14</b>	<b>References .....</b>	<b>35</b>

**List of abbreviations**

G-BA	Gemeinsamer Bundesausschuss (Federal Joint Committee)
SHI	Statutory Health Insurance
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care)
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
MTC	Mixed Treatment Comparison
NYHA	New York Heart Association
RCT	Randomized Controlled Trial
ROC	Receiver Operating Characteristic

# 1 Background

High quality data on the health effects of a therapy are obtained from clinical trials, ideally randomized controlled trials (RCTs). These results are entered on the benefit side when assessing the relation of benefits to costs of an intervention. Economic data are not regularly collected in clinical trials. If this is done, however, these data alone are often not sufficient for a full and substantiated depiction of the costs of a health technology. Clinical trials seldom provide information on the long-term economic consequences associated with the introduction of a new technology. In addition, they do not always adequately and comprehensively reflect all cost aspects relevant to the German health care setting [1,2]. Moreover, protocol-induced resource consumption in clinical trials may bias cost estimation. For these reasons, the modelling of the economic effects of a health technology is an essential component in health economic evaluation.

If the medical benefit assessment carried out before the economic evaluation is followed by modelling, modelling techniques can be introduced for the estimation of both costs and health consequences. This could be the case if the health technologies to be compared have a simultaneous effect on several aspects of health, which then have to be combined in a comprehensive score (e.g. quality-adjusted life years or other aggregating measures), or the medical benefit has to be transformed in order to be plotted on a cardinal scale. In addition, modelling permits predictions to be made as to how the effects observed in clinical trials would appear under different conditions and time perspectives. Prognostic adjustments can then be made. Given that any prognosis is associated with large uncertainty, it must be ensured that no “new” benefit is created, e.g. by the inclusion of additional outcomes or by reversing the results of included clinical trials.

## 2 Definition of the term “model”

Several definitions have been offered for the term “model” as applied in the context of health care. Models are analytic tools used to understand real-world systems, estimate outcomes for a given set of inputs and examine the effects of changes to the system being modelled [3]. In effect, any evaluation that extends beyond direct application of observed data can be considered a model [4]. To a certain extent even the statistical analysis of data involves aspects of modelling, for example in the case of meta-analyses of clinical trials. It is understood that models cannot represent reality perfectly: they are based on a reduced set of components and require simplifying assumptions [5]. Nevertheless, it is crucial that the model be valid in the sense that it sufficiently reflects the system it represents.

ISPOR (the International Society for Pharmacoeconomics and Outcomes Research) defines modelling as “an analytic methodology that accounts for events over time and across populations, that is based on data drawn from primary and/or secondary sources and whose purpose is to estimate the effects of an intervention on valued health consequences and costs” [6]. A recent publication on modelling techniques adopts a similar definition: “a formal quantified comparison of health technologies synthesizing sources of evidence on costs and benefits” [7]. Both definitions are compatible with the requirements of health economic evaluations conducted by the Institute for Quality and Efficiency in Health Care (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, IQWiG).

### **3 Aim of health economic modelling**

The aim of health economic modelling is to generate expected values for the clinical and economic effects of therapeutic alternatives. Modelling within a health economic evaluation conducted by IQWiG hereby provides information regarding the efficiency of an intervention, which is needed to define a maximum reimbursable price [8]. This working paper gives recommendations on how to develop a model, select a modelling technique, account for uncertainty and variability, and how to validate the model. In addition, the way in which modelling studies and their underlying assumptions have to be documented and described is depicted.

Although in principle IQWiG can be requested to assess various types of health technologies (e.g. preventive, diagnostic and therapeutic procedures), the main focus of this working paper is on technical aspects in modelling therapeutic procedures.

## 4 Process for developing a model

The process to follow in model development and in selecting the modelling technique, as well as validation, analysis and reporting, is outlined in this working paper. Briefly, the recommended steps involved in model development are as follows:

1. Define the research question(s) that the model will be expected to address. This includes:
  - a. What therapeutic options are assessed?
  - b. What population(s) is (are) relevant to the analysis?
  - c. What are the relevant cost categories associated with the therapeutic area and the interventions evaluated?
2. Develop an influence diagram, which graphically represents the factors that might have an influence on the research question(s) to be modelled and which shows possible associations between these factors [6].
3. Closely related to the development of the influence diagram is the specification of the model concept that serves as the design plan for the model. The model concept should illustrate how the health effects of the evaluated therapeutic options are linked to concrete health outcomes and resource use, and how they are projected over the time horizon of the model. In addition, the concept should explain how interconnections in the influence diagram will be dealt with, and provide justification for these choices. The model concept will determine the choice of modelling technique appropriate to answer the research question(s) (and not the reverse – i.e. the modelling technique should not be established first).
4. Data to populate the model should be systematically searched for only after the model concept has been established. It is not good practice to design the model around the available data, instead the model should be derived from the influence diagram and the model concept. If the systematic data search does not provide cost data for specific aspects of care, supplementary targeted data collection or primary data acquisition should be considered. This may be done through observational studies or by a targeted analysis/interpretation of existing registries or databases that originally collected data for other purposes (see Section 8.2.4)
5. Assessment, and possibly collection of additional data is followed by development of the functional relationships that form the core of the model. The type of relationships to be quantified will be determined by data availability and modelling technique selected. It is essential that these mathematical relationships be fully documented. In the case of *de novo* equations (e.g. derived from statistical regression modelling), documentation of their development is required.

6. The modelling technique that will be used to structure the model should be chosen concurrently with development of the functional relationships, as these activities are closely interrelated. The main criterion for selecting a technique should be whether it is suitable for implementing the application of the model concept as designed [9].
7. Implementation and programming of the model should only proceed upon validation of the model concept, selection of data, development of functional relations to be used and choice of modelling technique.
8. During implementation, technical validation should already be in progress. This is completed, along with face validation, upon finalization of the programming.
9. After the validation of a model is completed, an analysis of the relevant therapeutic options and target populations can be performed. These analyses must cover the most likely scenarios and address uncertainty in the input parameters or in the model structure [10].
10. The report on modelling should encompass all the procedural steps described above and summarize the results thereof. Reports must be accompanied by a fully functional and accessible electronic model. In order to ensure transparency and encourage confidence in its integrity, the model must be made publicly available.

Developing a valid model is a substantial task and, as such, the use of previously validated models is encouraged (see Section 12). IQWiG will retain a bank of models developed under its purview, which can be used for further economic evaluations. Nevertheless, modifications and updating of existing models will be required.

## 5 Influence diagrams

### 5.1 Definition and purpose

Influence diagrams are being increasingly used for the development of economic models [11]. An influence diagram graphically depicts the essential relationships of the model. It displays the factors that have or might have an influence on (the) research question(s) to be modelled. The diagram thus enables all interested parties to understand the modeller's idea and enables them to submit recommendations for changes without having an understanding of the actual technical aspects.

Despite its name the influence diagram does not per se show causal associations; rather it reveals the relationship between factors without necessarily having to take into account any stochastic interdependence among them.

### 5.2 Components

The influence diagram must include all important features of the disease being modelled and the relationships among these, including:

- Patient characteristics
- Pathophysiologic processes and clinical course of the disease
- Treatment and disease management
- Effects of an intervention

These components are typically depicted by oval shapes, and the relationships between them by using arrow connectors with the arrowhead indicating the direction of the relationship (without necessarily indicating causality).

An explanation of the components and connections in the influence diagram should be provided in a written document accompanying the diagram. In this document, each component (i.e. oval shape) and connection is explained. The rationale for omitting connections that might have been included should also be given.

## 6 Model concept

### 6.1 Definition and purpose

The model concept depicts accurately and in detail the analytic structure and the approach of the model. It is based on the influence diagram and provides the intended design in much greater depth.

Defining the model concept is a critical step in model development and should therefore be well documented and described. The model concept should increase the understanding and substantiate the validity of the later model. Its purpose is to clarify how the disease area being modelled and the effects of the interventions are understood and assessed. All assumptions about relationships between model components should be explicitly mentioned. As even the most sophisticated models are simplifications of reality, with required assumptions and limitations on what can be included [10,12], it is only by clearly specifying the concept that the model's validity and utility can be properly understood.

### 6.2 Components

Specifying the model concept should take into consideration the following elements:

- What clinical events are to be included and what influences their occurrence?
  - Can these events recur? What influences the risk of recurrence?
  - Does the occurrence of an event influence the likelihood or timing of subsequent events?
  - Does the severity of events needs to be evaluated?
- Which conditions are to be considered?
  - Do the conditions sufficiently describe the clinical spectrum of the disease?
  - Are they sufficient to differentiate patient groups according to health status, health-related quality of life, and costs?
  - Do they sufficiently capture the patient's history and experience?
- How does the patient-relevant, additional therapeutic benefit of the intervention(s), as determined by IQWiG, influence the natural course of the disease?
  - Are events delayed or averted altogether and/or is their severity reduced?

- What are the underlying mechanisms and assumptions which determine how the intervention affects the natural course of the disease?
- What assumptions are made on long-term effectiveness of the interventions, and on consequences of treatment delays or discontinuation?
- How does the behaviour of patients/providers influence outcomes?
  - How important is the patients'/medical providers' adherence to a therapeutic regimen and how does this influence future medical decisions?
  - Is interaction between patients relevant, as it would be, for example, in the case of acute infectious diseases?
- How are costs influenced by patient management, patient history, patient behaviour, and the timing and severity of complications?

### **6.3 Validation**

Models should be “consistent both with a coherent theory of the health condition being modelled and with available evidence regarding causal linkages between variables” [6]. Thus, finalization of the core model concept cannot proceed without validation.

In order to validate a model concept, it is important to include full documentation of the information used to define the model concept and the hypothesized linkages.

It is also helpful to compare the concept with that of relevant existing models (i.e. cross-model validation). This comparison should include a description and justification of any deviations from concepts normally applied in the disease and therapeutic areas in question.

Given the uneven quality and the lack of a “gold standard” for the components of a “good model”, concordance with published models can be insufficient [13]. It is crucial that models must make sense to people (e.g. acknowledged experts, clinicians, patients and their families) with knowledge of the disease [14]. Validation of the model concept, therefore, must involve appropriate experts including individuals familiar with the therapeutic areas being modelled.

## 7 Features of the model

To be considered appropriate for IQWiG purposes, models for the assessment of benefits and costs of a technology as compared to therapy alternatives must conform to a minimum set of requirements. This section outlines those requirements, though it should be noted that the quality of models is not based on features alone, but also on peer reviewing, which is possible by providing fully accessible models.

### 7.1 Appropriate depth

The models developed by IQWiG have to show the costs and benefits generated by a specific health technology in comparison to alternative technologies used for the indication under consideration. This requires the following information:

- Precise definition and differentiation of the additional health benefit that is realized with the intervention.
- If possible, all relevant aspects of the disease and its treatment should be captured. This will most likely have an impact on the cost or benefit components of the model, for example in the following areas:
  - Demography
  - Epidemiology
  - Specific treatment patterns
  - Adherence to treatment.

Where sufficient data are not available to incorporate these features, it is essential to explore the possible impact on cost-benefit estimates of reasonable ranges of values for the missing information.

### 7.2 Perspective

The primary perspective of the analysis should be that of the community of Statutory Health Insurance (SHI) insureds. This is not identical to the perspective of the SHI itself, as it includes resources of patients.

If possible, cost results should be reported in aggregated as well as in completely disaggregated form. A comprehensive description of cost components is provided in the working paper “Cost Estimation”.

### **7.3 Time horizon**

The cost-effectiveness of interventions should be separately modelled for at least two treatment scenarios:

- First scenario: treatment lasts as long as it did in the trials (usually RCTs) based on which treatment benefits were assessed. This permits a parallel validation of the model through existing clinical trial data.
- Second scenario: treatment duration extends beyond the duration of existing clinical trial data, if this is relevant to the decision maker.

### **7.4 Discounting**

All costs and benefits used in the model must be reported at their properly discounted value. The sensitivity analysis should cover a broad range of discount rates to investigate the effect of discounting on the intervention's net cost and health effects. Discount rates will be set by IQWiG and updated if necessary. Detailed information on discounting costs and benefits is provided in the main methods paper [15] as well as in the working paper "Cost Estimation" [16].

## 8 Assessment of data

The available data should be systematically collated, and their quality and relevance as a model input should be investigated with particular consideration of its risk functions. Evidence hierarchies, which are used to classify the validity of evidence from different types of studies, are also applied in health economic studies [17]. This is a reasonable approach with respect to evidence on data on benefit [18], but not always with respect to cost data. Randomized controlled clinical trials often do not comprehensively capture the specific cost-relevant aspects of care [17]. For cost information, controlled prospective studies are superior to retrospective studies. However, a retrospectively created database cohort or a case control approach may also be helpful in the collection of cost data. Given that non-randomized studies do not per se adequately reflect the cost-relevant aspects of health care processes, they likewise have to be cautiously assessed with regard to the appropriateness of the information provided.

Typically, data will be drawn from several sources, including trials on treatment effects, cohort studies that investigate specific parameters and risk factors related to the natural course of the disease and the associated life expectancy, trans-sectional surveys collecting quality of life data, registry data on resource use and costs as well as automated databases or compiled statistics. Assumptions based on expert opinion should be avoided, as they are rarely accurate enough to be used when other sources of evidence are not available. An important criterion for the studies is the transferability of results to the German context. Due to the differences between systems, transferring cost data from other health systems is seldom possible and, if done, then only under stringent conditions [3,19].

The rationale behind the selection of specific data sources should be documented, especially for the key parameters which may have a substantial impact on the model results. Details of the data processing methods which are used to develop the required inputs, functions or distributions of parameters must be provided [13,20,21].

### 8.1 Objectives

In order to verify the available data, the following steps have to be taken:

- review its quality and assess its relevance to the specified model objectives,
- identify any data gaps for which assumptions will be appropriate, and
- apply a suitable method to process the data in order to develop reliable effect estimates.

## 8.2 Potential sources

### 8.2.1 Clinical trials

IQWiG assesses the benefit of an intervention based on patient-relevant outcomes from clinical trials, especially with respect to the outcomes mortality, morbidity and health-related quality of life as well as valid surrogate parameters. Criteria, with which valid surrogates must comply, are described in IQWiG's General Methods [18]. The patient-relevant beneficial effects of preventive, diagnostic and therapeutic interventions are determined in IQWiG's benefit assessments based on data from RCTs.

The model may use inputs that are based on individual trials or on pooled trial datasets like a meta-analysis of reported study results. At the time of review of reimbursement for a new product, it is often the case that only relatively small and short-term efficacy trials have been completed. The duration of these efficacy trials may be insufficient to detect significant differences on the cost side. This tends to be a major limitation of economic evaluations because the latter need to reflect treatment practices in health care that may extend to broader time horizons than the ones reflected in efficacy trials [22-24].

Meta-analyses of trials are often used to mathematically summarize the treatment effects of individual trials. Head-to-head comparisons of each treatment are not always available, so that indirect comparisons have to be included to ensure that all the pertinent treatments are considered. Appropriate methods should be used to derive these indirect estimates of treatment effects [25-27]. The mixed treatment comparison (MTC) meta-analysis is preferably used in these cases. In MTC meta-analysis, direct and indirect evidence are combined to estimate treatment effects [28-34].

As data on cost-relevant aspects of health care are often not comprehensively captured in clinical trials, the use of data from these clinical trials to estimate costs may be misleading. On the other hand, protocol-induced costs can be incurred, which are not to be included in a health economic evaluation [1,2]. If data are extracted from multicentre or international clinical studies, it should be verified that the treatment patterns observed in these trials adequately reflect context-specific resource consumption in Germany. Country-specific differences, e.g. with respect to the choice of comparators or with respect to the compliance of patients/providers, could have a confounding effect on any cost estimation [22-24]. In addition, epidemiological parameters that influence cost estimation can considerably vary between countries [19]. Therefore, any data extracted from such studies cannot be transferred to the German context without further analysis [3,35-37].

Given these data limitations or if cost data are unavailable from clinical trials, the use of alternative data sources for the estimation of resource use, e.g. from observational studies, may be considered. Due to their importance for model results, all data sources that are used to estimate resource consumption should be described and well justified.

All steps in the transformation of data from the benefit assessment to a form that is fed into the health economic model should also be described and any adaptations in the data should be made transparent.

### **8.2.2 Epidemiological studies**

Data from observational studies are sometimes necessary to enlarge or extrapolate clinical trial data and to facilitate modelling beyond the time horizon of the underlying RCT(s). For many diseases it takes years or decades until the final health outcomes become manifest. In such cases, it is common to assess intermediate clinical outcomes that serve as surrogates for the final patient-relevant outcomes, such as long-term morbidity and mortality [8].

The interpretation of predictive functions and the validity of transferring established risk functions to the modelled population, as well as any adjustments made, have to be explained and justified. Reasons for the use of data from observational studies should be stated. Details must be provided on the statistical methods applied and on any analyses which are performed to develop appropriate estimates for incorporation into the model. The validity of the assumption that these functions or estimates can be applied to the German population and thus used for the model needs to be addressed and any modifications explained.

### **8.2.3 Automatically generated databases**

Automated claims databases can provide details of health care resource use by a large number of patients in real-life practice. They are thus a useful data source when the diagnostic codes employed by the database are sufficient to specify the relevant patient population. This may be problematic if the definition of the population includes clinical findings, laboratory results, patient characteristics other than age and gender or other parameters that are not typically encoded in such databases. Data on claims for medications can also provide information on real-world treatment paths, switching, adherence to treatments (at least regarding obtaining a prescription if not regarding the actual utilization of medications), and persistence, and thus can be utilized in the cost determination.

As many factors can impact on prescribing practices, including local reimbursement conditions, product launch dates, and even delays in publishing data, the relevance of a given database to a specific model has to be assessed. Although a given database may be useful for quantifying resource use patterns for a given condition, caution must be exercised in comparative analyses, as there may be severe confounding when patients are not randomly assigned to treatments [20].

### **8.2.4 Registries**

Patient registries are observational studies that systematically collect a reduced set of data from large numbers of patients managed in the routine care setting. As the data are usually

collected in the regular care setting, they provide a good representation of current clinical practice.

A well-designed registry can be a valuable source of data for a model. It can support analyses to develop many of the required predictive functions describing the association between clinical events and resource use, quality of care, work days missed, etc.

Assessments of the value of data from a registry should include evaluating its relevance, taking account of the registry objectives, design, patient population(s), and conduct, in order to confirm the quality and completeness of the data collected (for example, by checking for the proportion of missing data on important variables and understanding how these were handled and reported) [38]. The analyses undertaken to develop functions and estimates that are in a form appropriate for inclusion in the model must be detailed. Additionally, appropriate access to the registry must be provided in terms of quality control.

### **8.2.5 Compiled statistics**

Compiled statistics are health statistics compiled by governments from various sources such as a census or survey. These sources might provide relevant data on demographics (e.g. age, sex) as well as health-relevant behaviour and risk factors (e.g. weight, smoker/non-smoker) [39]. Compiled statistics are often used to estimate mortality for causes other than those of interest in the model (e.g. the Human Mortality Database [40]). Any sources of compiled data used in the model must be provided and assessed in terms of quality and relevance to ensure that they are appropriate, reasonably up-to-date, and complete.

### **8.2.6 Expert opinion**

Expert opinion, focus groups and consensus panels may be used in modelling studies, but only to establish the model structure and to support the assumptions on which the influence diagram and the model concept are based. If no data are available, a pragmatic approach may be taken in health economic studies and evidence replaced by expert opinions [13,21,41]. However, this can be considered a methodological flaw [42]. This source of information is almost never appropriate for actual input values and for the development of functional relations [20,43-45].

Expert opinion may be considered valuable if the complete process of patient management is to be documented. Experts can facilitate the transfer of data on resource use between countries, if equivalent data are not available in their own country. The justification for using expert opinions and a description of how they were gathered must be included in the model report [43]. The variability in the opinions obtained must be reflected [13,43,45,46].

## 9 Functional relations

### 9.1 Definition

A functional relation is an equation relating a response variable to its determinants, that is, to variables that are correlated or associated with the response. The covariates (determinants) will typically include patient characteristics and time. These functional relations are derived from the data sources used to support the model by employing appropriate statistical techniques. Alternatively, published equations may be used if they are suitable for the population modelled. This section describes different types of relations and statistical techniques suitable for their derivation.

### 9.2 Purpose

Functional relations should be specified to reflect the relationships depicted in the influence diagram of the model. In other words, these equations should be used to ensure that parameters in the economic model that were felt to be associated are appropriately linked, so that a change in one parameter is reflected by a change in the other (here termed "effect" without necessarily implying causality). Ignoring relationships between parameters, or equivalently, omitting important factors from the functional relations specified in the model, leads to an averaging of the effect (on the response variable) over levels or values of the omitted factors. Incorporating correlates of effect and costs parameters yields more precise estimates, thus reducing the uncertainty of the economic results [47-49].

### 9.3 General concepts

Functional relations are particularly important in simulation-based economic evaluations as they control the simulation. Multiple regression techniques and other methods, e.g. survival analysis, can be used to estimate the parameters for these functional relations [50-55].

Functional relationships should be specified to relate benefit and cost parameters to patients' characteristics, treatment, time and other factors that may be of importance. At its simplest, the effect parameters would be related to the single determinants of interest only: thus, when addressing the intervention, these would yield "average" responses for the intervention and for any alternative under consideration. Including patient characteristics that are associated with the response variables, on the other hand, allows patient-specific predictions of changes in the affected parameter. This framework also allows the integration of subgroup effects through the inclusion of interaction terms. Resource use can be modelled similarly to capture variations among patient types; for instance, older patients or those with more severe illness may incur greater costs. However, the statistical properties of resource use data (right-skewed distribution of costs data) need to be handled very carefully when modelling [56].

If input variables are associated, the functional relation between these variables should be integrated into the model. For example, large reductions in hospitalization can be accompanied by large increases in home care services and vice versa. In addition, risk factors correlate frequently. For instance, weight may be associated with the patients' sex as severity of illness may be with patient characteristics. To capture these correlations in modelling, the Cholesky decomposition method (for multivariate normal distributions) or bootstrapping methods that account for correlations between variables can be applied [50].

Continuous variables can mostly be modelled with linear regression techniques. However, this is not always the case for cost data, which tend to have a (highly) skewed distribution. Log transformations or alternative models based on gamma distributions should be considered [47,57,58]. When the asymmetry in costs is caused by an excess number of patients with no costs, a two-step approach is appropriate, in which the probability of incurring any cost at all is modelled first, followed by analysis of the magnitude of the costs for patients incurring expenses [59].

If data are collected at different locations/study centres, it is important to analyse the influence of the location on both the effect and cost parameters [47,58,60]. Potential bias must be controlled by using appropriate techniques [58,61-63].

Evidence supporting the validity of the functional relations used in the economic model should be presented. This should include a description of all the assumptions and decisions made in developing the equations. To show that data sources were adequate to be used in the conducted analyses, sufficient information should also be provided about them.

When developing a prediction equation various strategies may be used to select variables. These include automated selection procedures (forward, backward, stepwise selection), or manual selection of variables based on statistical significance, or variables based on prior clinical knowledge about the relationships being examined. In accordance with its methods (Allgemeine Methoden 1.0) IQWiG recommends backward selection [18]. As complete absence of clinical knowledge about the relationships is extremely rare, it is recommended that prior clinical knowledge is taken into account in any model building procedure. In any case, the approach used to select the variables for each equation should be described and justified.

Explicit information should also be included about the criteria used in selecting the appropriate form of equations. For instance, in the case of a linear equation the underlying assumption can be that the criteria may be to maximize explained variability. In logistic and survival analyses, predictive functions may be evaluated in terms of their discriminating power. Measures such as the c-index or receiver operating characteristic (ROC) curves can be used [51].

The assumptions underlying the functional relations should be presented, along with approaches used to ensure that the assumptions are "valid". The goodness-of-fit of chosen

regression models should be investigated and described by means of common procedures [51,64]. When developing causal functions, the assumptions on the existence and direction of causal associations should be explicitly stated using causal graphs [65].

## 10 Modelling techniques

A number of modelling techniques such as decision trees, Markov state-transition models, discrete event simulation (DES), agent-based simulation, transmission models, and others are available and have been applied in health economic evaluations [4,7,12,66-70].

The description of modelling techniques given in the following sections is by no means exhaustive, but covers the most commonly used techniques in health economics. For additional literature, readers are referred to Brennan's taxonomy of model structures [7], Koopman's examination of infectious disease models [71], and Stahl's overview and guide on modelling methods [69].

### 10.1 Choice of modelling technique

It is vital that no specific modelling technique is chosen in advance. Hence, IQWiG has no *a priori* preference for a particular modelling technique. The choice of appropriate modelling technique depends on the research question posed by the Federal Joint Committee (G-BA) and commissioned to IQWiG. The properties of the evaluated technology as well as the characteristics of the disease and the setting are further criteria for the choice of a modelling technique. The proper process to develop an efficient model that addresses the relevant problems is to specify and fully understand the model concept first. Also, the available data and functional relations have to be assessed before a modelling technique can be chosen. The modelling technique selected by IQWiG for a specific research project will be described in the respective preliminary report plans and put forward for discussion.

Ultimately, the choice of a technique will be influenced by a number of factors, but the guiding principle remains that the economic model should adequately answer the research questions being posed. Simplicity of design for the sake of minimizing process and analysis time, or to make the model understandable to lay persons, is not recommended. Equally unacceptable is the selection of a modelling technique purely on the grounds of familiarity. With adequate documentation, even "complex" models can be fully transparent. Increased computer processing capabilities and more up-to-date variance-reduction techniques ensure that time and computing requirements to run complex models are decreasing [9,70,72,73]. If the choice of a technique leads to modification of the model concept, the choice might be wrong [5,74].

### 10.2 Aspects in the choice of modelling technique

Although many characteristics of models could be considered in defining the choices to be made [7], there are two aspects that are particularly important for health economic evaluations. The first aspect is the level at which the population is modelled (cohort versus individual level). The second aspect is the form in which the course of a disease is presented (state transition simulation versus event simulation).

### 10.2.1 Cohort versus individual models

Broadly speaking, models can be divided into those that address individual patients and those that do not distinguish individuals, instead using a cohort approach. Cohort models aggregate the individuals into a group which becomes the unit of analysis. Over time this group “breaks up” into pre-defined subgroups according to the events being modelled. The initial group is defined by a single set of characteristics (e.g. age, gender, disease). The subsequent subgroups are also defined based on their own set of characteristics (e.g. new stage of disease). Individual models consider the experience of each patient individually, even if they report results at the level of the entire population. Each individual has unique characteristics, on the basis of which their individual course is modelled. Although these individual models typically require more data, computing power, and running times [70,72] this does not necessarily mean that they are more difficult to fill with data and more difficult to understand.

Cohort models are the most common technique used in health economics today [12,75]. The individual simulation (often termed “micro-simulation”), once nearly unheard of in health economic evaluations, has become more prominent in recent years [70,72]. Even though there is resistance to its use [76], it is widely accepted that in many situations individual-level simulation can provide more flexibility and accuracy. Also, micro-simulation models can be used for a wider range of health care questions. On the other hand, if the individual patient characteristics and their heterogeneity in the target population are not of great relevance to the decision problem, cohort simulation may be sufficient to answer the research question [4,7,9,50,68,70,72,77,78].

For IQWiG health economic evaluations, the choice of cohort simulation versus individual-level simulation must be carefully justified. If a cohort technique is employed, it must be demonstrated that the choice of a cohort approach did not inappropriately modify the model concept and that it does not introduce distortions that represent significant departures from reality. If an individual-level simulation is employed, it must be demonstrated that appropriate distributions and covariance matrices have been used for the individual model parameters and that the assumptions on these distributions are not unrealistic simplifications of reality.

### 10.2.2 Event-based versus state transition-based simulation

The most commonly used technique in modelling in general is one that considers the events that can happen in a discrete event simulation [79] or related techniques [80]. In health economic evaluations, however, the state-transition concept (Markov modelling) is far more prevalent [75]. Markov models were introduced in health care evaluation in the early 1980s [81] and have a long history of use in health care decision making [82].

In a Markov model, the population moves through several (mutually exclusive) health states in specified time intervals (Markov cycles). During this process, health outcomes and costs are accumulated and can be compared between the interventions of interest [83]. The

technique of Markov modelling is described in detail in section 10.3.3 including its most relevant advantages and disadvantages.

In contrast to the state transition-based models, event-based simulation [79] conceptualizes reality in terms of the events that can occur. This type of simulation allows for a representation of the natural course of a disease, treatment patterns, compliance with treatments, and other relevant factors. As time is both explicit and continuous, and events to be modelled can be anything that happens in reality and occurring in any order with no *a priori* restrictions, the experience of patients can be captured very accurately. The technique of discrete event simulation is described in section 10.3.2 in more detail.

Event-based models are powerful tools, which are currently still under-used in health economics. Given the flexibility and realism of event-based simulations, it is recommended that these should be given equal consideration for use in IQWiG's health economic evaluations. If an individual, for instance, were to experience a stroke event, an admission to hospital event, and a death event within a single day, and this sequence were to be captured by the model, an event-based approach might be the best choice, as it allows for more accurate application of costs [72].

The choice in favour of the state-based instead of the event-based modelling approach must be justified and it should be shown that the choice of a particular technique does not inappropriately modify the model concept or introduce unacceptable distortions of reality.

## **10.3 Techniques**

### **10.3.1 Decision trees**

A decision tree is a visual representation of all the investigated technology options and the consequences that may follow each one [84]. Each intervention is followed by branches representing the possible consequences with their respective probabilities. Probabilities may depend not only on the intervention but also on patient characteristics. At the end of the tree, each path leads to an outcome, for example, the alternative outcomes of a clinical trial. For each of the therapeutic alternatives, the expected value of the clinical and economic consequences can be calculated as a weighted average of all possible consequences, applying the path probabilities as weights. Decision trees work well for very simple decision problems, where there is limited repetition of relevant events and a limited and fixed time horizon is analysed. Decision trees can be calculated either with cohort simulation or with patient-level simulation.

### **10.3.2 Discrete event simulation**

In the area of health economics, the use of discrete event simulation has become more prominent in recent years [4,67,69,72,77,85]. Nevertheless, the basic approach to discrete

event simulation is well covered in the non-medical literature [86-94]. Due to the growing importance of this technique in health economic evaluations, a full description of it is provided in this section, and its advantages and disadvantages are highlighted compared to the state transition-based techniques.

Discrete event simulation operates at the level of individual entities which are the objects that will experience the events considered in the model. There can be many kinds of such objects in a discrete event simulation, and the entities may even change type during the simulation. In health economic models, a nearly universal type of entity is the patient, but the basic entity could also be defined at the level of organs in a more detailed model. Other common entities are health professionals, caregivers and even inanimate objects. Entities are created and assigned a type at any point in the simulation via an “entry event”. In health economic evaluations, it is customary to create all entities at the beginning of the simulation but this does not necessarily have to be the case. Entities can enter the simulation later due to incidence, births, immigration and other such mechanisms. Entities stay in the simulation (even if nothing is happening to them) until an “exit event” occurs at which point they leave the simulation. Although death is the most common type of “end event”, entities may leave for other reasons such as end of their time horizon, loss of the indication, emigration or similar.

Entities are individualized by assigning specific values to their attributes. All entities share a common set of attributes (e.g. all would have an “age” if that attribute was part of a given model) but individual entities are assigned their own unique value for each attribute. These individual characteristics are updated throughout the course of the simulation, due to the natural flow of time, or to events that the entity experiences (e.g. medical complications, medical consultations), or due to other events affecting the entire system (e.g. a change in reimbursement, introduction of a new process) or even to the experiences of other entities (e.g. death of a caregiver). Aspects of the history of each individual can be retained as attributes if they affect the occurrence of subsequent events and decisions about management, or other attributes and any other aspect modelled. These attribute values are accessible to the model and can be output for display or further analysis. Needless to say, heterogeneity in the population, treatment effect, and outcomes can be incorporated readily into the simulation as the attributes are individual.

In a discrete event simulation, time is explicitly modelled and it runs continuously as it does in reality. This means that the timing of events is unrestricted – they can even occur simultaneously (within the same day, for example) and competing risks can be properly applied. There is no need for “half-cycle” corrections or other procedures to reduce the error introduced by fixed time cycles. It also enables the analyst to appropriately implement any time dependencies that exist and to keep track of the timing of events, the duration of conditions and so on. Delays – both planned (e.g. one day in hospital for a procedure) or unplanned (e.g. waiting for a bed to be available) – can be incorporated.

A particular strength of discrete event simulations is that resources are also modelled explicitly. They are considered special events where a resource is used. These resources can be people (e.g. medical practitioners, nursing staff, other health professionals), goods (e.g. drugs, intravenous equipment), organs (e.g. liver transplants), space (e.g. hospital bed) or other aspects that are conceptualized as a resource. The resources have unit costs which can be specified per “use”, per hour, and so on. In addition, these costs can be time-specific (e.g. more expensive during out-of-hours); even the costs of inactive time can be specified. Resources also denote units of capacity such as the number of physicians within a region or the number of beds in an intensive care unit). Thus, a type of resource can become a “bottleneck”, creating queues as individuals who need the resource wait for it. If such resource limitations are a relevant feature of the decision problem in an IQWiG evaluation, it is recommended to choose a discrete event simulation rather than a Markov model.

Another advantage of discrete event simulation is that individuals may interact with each other [4,7,89]. For example, discrete event simulation can be integrated in a model, with individual entities “competing” for those resources, and queuing up to use a resource. In health care, where constrained resources are ubiquitous, assessing how new interventions might free up limited resources can be critical. Other examples include changes in behaviour. For example, as medical practitioners gain experience with a health technology and see more patients, their efficiency and subsequent treatment patterns for new patients may be influenced by their experience with previous patients. Finally, in the area of infectious diseases, where transmission of infections is a direct function of the infectious status of the individuals in the community, the ability to specify interactions and infection rules between entities is highly pertinent (see Section 10.3.5).

Discrete event simulations as well as individual-level Markov models (Section 10.3.3) use different types of functional relationships, which are mathematically captured. A discrete event simulation uses distributions like the exponential, Weibull, beta or gamma distributions, as well as others, in order to represent variability e.g. in inputs and event times. Together, these create the stochastic behaviour of the model. Random numbers must be generated by advanced algorithms that ensure sufficient “randomness” in the sequence.

Given the stochastic nature of discrete event simulation, any single run of the model may yield results that are far from typical. Thus, it is important to replicate the runs many times. Each replication uses a different set of random numbers and thus all the processes that use random numbers, including the creation of the population and all of its experiences are unique. This yields a different set of results for each replication. Over many such replications, the result values will tend towards their “average”, providing a more stable estimate of the results. A common question is “How many replications are needed?” There is no single answer to this question as it depends on the variability in the model, the patients, and so on. The best way to check this is empirical: examine the standard deviations of the critical results between replications. When the width is acceptable, there have been enough replications.

While it is possible to program discrete event simulations to act like cohort models – a temptation when data are limited – the technique’s flexibility allows for the simultaneous specification of multiple levels of complexity within the same model. In cases where data are scarce, the same model can be run with a simple set of inputs (e.g. by using the average treatment effect instead of sampling from a distribution of treatment effects), but also in more sophisticated scenarios. This allows the analyst to address how the results might change if more detailed data were available to populate the model (e.g. what would happen if the treatment effect is allowed to vary at the level of the individual).

Modern software allows for a very transparent representation of the relationships in the model, for example, possible patient flows through the simulation. Software packages such as Arena® [89] allow for programming using flowchart methods very similar to those clinicians are already familiar with. They provide a visual representation of patient flows and possible outcomes. As discrete event simulations can produce a wide range of intermediate and final results, assessing the outcomes at numerous levels is possible. This is useful for validation purposes, and also provides a much more comprehensive set of results that can address the concerns of different decision makers and analysts.

### **10.3.3 Markov models**

Markov models are suitable for modelling prevention, diagnosis, and treatment of chronic diseases, where (1) parameters are time-dependent (e.g. restenosis after coronary artery intervention, relapse after cancer treatment), (2) time-to-event is important (e.g. the time to an event such as stroke, cancer relapse etc. matters), and (3) repeated events may occur (e.g. restenosis, second myocardial infarction). Likewise, numerous guidelines on how to design Markov models have already been published [75,82,95]. Some of the strengths and weaknesses of Markov models are highlighted again in this section. Extending this model technique to a simulation at the level of an individual is described.

In a state transition Markov model, either the entire cohort starts at time zero in an initial disease state or the cohort can be distributed between different Markov states. The population in each state is defined by a common set of characteristics which describe that segment of the population (e.g. age, sex, disease severity) with no distinctions among individuals [72,75]. These characteristics can affect the likelihood of moving to other states in the next cycle (e.g. proportion with disease who die depends on age) and the “valuation” of the state (e.g. costs accumulated during each cycle depend on disease severity) [95]. The characteristics of each state can also change over time. In each model cycle, the distribution of patients regarding the various states is recalculated based on the relevant transition probabilities. Running this analysis for many cycles creates a profile, which determines how many patients are in a specific state at a particular point in time [95,96].

State transition models are based on several simplifying assumptions and have limitations. Patients (whether as fractions of a cohort or as individuals) are described in terms of the states

in which they find themselves in a particular cycle [82]. These discrete “states” are mutually exclusive (no part of the population can belong to two states in the same cycle), are assumed to be homogeneous and are complete (the sum of cohort fractions in all states added together must be 100 % in each cycle). There is no exit from absorbing states (e.g. death, amputations): the fraction of the cohorts that reach this state remains there for the rest of the modelling period. Therefore, a sufficient number of states must be chosen to represent the full spectrum of the disease. Moreover, if a person acquires two or more properties (e.g. history of myocardial infarction and stroke), Markov states must be created for such combination states. In addition, there is exactly one transition between states per person and cycle. Markov models handle time by specifying fixed cycles over which the transition probabilities are applied, with any segment of the population allowed only one transition per cycle. In order to properly handle competing risks, additional techniques are required [72]. The analyst must choose a sufficiently small cycle length to mimic the continuous time flow of reality.

A strong assumption of Markov models is that does not incorporate a memory of previous states (the “Markovian assumption”) [70,75,95]. In other words, persons who move from one state to another acquire all the characteristics of the actual state and lose all characteristics of prior ones. The likelihood of moving from one state to another depends solely on the current health state. It is not dependent on prior experience nor on the duration of time in a state. As the transition probability between different disease stages is often determined by factors influenced by earlier stages, the assumptions stated in the previous sentence do not represent reality. This means that relevant experience such as prior events must be included in the description of the states. For example, “tunnel states” must be created to reflect the time a patient has already spent in a specific disease stage. This can easily lead to an explosion of the number of states thereby making the Markov model difficult to handle.

The profile of a state may not take into account all the relevant characteristics needed to model transition. This can lead to substantial distortions (e.g. with one-year cycles, the mean age is typically increased by one year even though it is the older people who may be transitioning out due to death and disease). This aspect of cohort Markov models can also be overcome by accounting for the heterogeneity in terms of additional health states that can be reasonably assumed to be homogeneous. For multiple patient characteristics, however, this can lead to a rapidly expanding set of states that soon becomes unfeasible. In both cases described in the section above, individual-level simulation of Markov models can be used to overcome the problem).

The simplest Markov models are chains that assume that the transition probabilities are constant over the entire modelling period [82]. Markov chains are a very attractive way to model because they can be solved numerically [97], by inverting the related matrices (i.e. it is not necessary to run the model at all). This property of constant transition probabilities is almost never realistic in health care, however. For example, background mortality rates of individuals change with increasing age. As decision-analytical software allows iterative solutions to Markov models which are not associated with the technical complexity of a

matrix algebra solution, health economic evaluations tend to use Markov processes that allow the transition probabilities to vary over time (e.g. for the risk of death to increase as the model progresses through time) [81]. Another aspect of Markov models relating to transition probabilities is the Markovian assumption of all states being "memoryless", already hinted at above in Section 10.1.2. In other words, the model has no memory of what has happened to the patient before. Therefore, a sufficient set of Markov states that captures the important patient histories must be chosen to provide a realistic model of the disease.

In a pure Markov model, interactions between individuals are not considered [4,7,67]. Meanwhile, some software has tried to relax this restriction by allowing "external" parameters such as disease prevalence to be used and updated in the model. This feature can be used to incorporate some transmission features in the Markov framework.

Individual-level Markov models are often referred to as Monte Carlo simulations because of the techniques used to sample probability distributions. In these models, individuals instead of a cohort are modelled. This allows each individual's own historical experiences to be taken into account. Monte-Carlo-simulations hereby overcome the restrictive (and often unreal) Markovian assumption of "memorylessness" and enable modelling while taking into account the entire clinical history of patients and the variability in their individual characteristics [82]. In this way, many of the limitations of cohort models can be overcome [7,75,82]. In these models each individual has their individual characteristics, historical experiences can be retained, and heterogeneity in treatment effects and outcomes can be accounted for. Unlike cohort models, Monte Carlo simulations are stochastic, and all elements of the simulation can be specified using probability distributions. Patient flow through the model is determined by sampling from these probability distributions. A difficulty arises when these models cycle one individual at a time applying all the transition probabilities each cycle even if the individual experiences no transition in the majority of cycles. This slows processing time substantially.

Similar to discrete event simulations, individual-level Markov simulations can use equations of many types to control the modelling process (e.g. linear functions, logistic regressions, hazard functions, Cox proportional hazards, etc.).

Markov models can be modified further in various ways to overcome the limitations of a pure Markov model. For example, "tunnel states", "half-cycle corrections", multiple sub-states, external macros, and other features of available software are routinely used in practice. However, these features counter one of the main appeals of Markov models – their simplicity.

As with discrete event simulation, new software has simplified the programming of Markov cohort and individual-level simulations, allowing for intuitive visual representations of the model processes and facilitating model transparency. In addition, there are pre-specified tools simplifying and supporting the visualization of the state distribution over time (known as the Markov trace), the carrying out of sensitivity analyses, presentation of uncertainties, and link to spreadsheet software and macro language.

### 10.3.4 Agent-based simulation

Agent-based simulation is a powerful further development of discrete event simulation, but even less commonly used as a modelling technique in health care [66,80,98-100]. The main difference compared with standard discrete event simulation is that in an agent-based model, the entities (called agents) are aware of each other and of their environment [80]. It is highly suitable for modelling infectious diseases where transmission of infection is important. If social behaviour plays a key role in influencing the outcomes, this modelling technique may also be used. While basic discrete event simulation can accomplish many of these things, agent-based simulation possesses the advantage that it takes interactions into account. It facilitates the programming process and allows for a more transparent presentation of how the model operates.

An agent, like an entity, will often represent patients, but may represent people who are at risk from the illness in question. Like an entity, agents can also be specified to represent other groups such as medical professionals, or even organizations and communities. Agents can modify their behaviour over the course of the simulation and influence the behaviour of other agents, and thus one of the critical components of an agent is its “capability [...] to make independent decisions [and] [...] to be active rather than purely passive” [80]. Agent-based simulation also allows for spatial relationships between agents to be specified.

As with discrete event simulation, new software has simplified the programming of agent-based simulations. Moreover, this software allows for intuitive visual representation of the model processes and thus facilitates transparency [66]. These visual tools are not only useful for the decision maker but also for the analyst when debugging.

Some of the scenarios where use of agent-based simulation may be considered [66] are as follows:

- behaviour of individuals is non-linear and possibly best described by discontinuous functions (e.g. thresholds),
- individual temporal correlations, adaptation and learning need to be represented,
- interactions between individuals are heterogeneous,
- network effects may be present,
- fluctuations are possible in the system, with a steady state based on average relationships being unrealistic.

### 10.3.5 Transmission models

In epidemiological and public health modelling of infectious diseases, transmission models are a much more common and accepted technique [71]. Transmission models [101] can be

constructed at various levels of complexity and can be either deterministic or stochastic [68,71]. In their simplest form, deterministic models are based on solutions to a series of differential equations which describe the infectious states (compartments) of the population (e.g. susceptible, infected, immune), the degree of infectiousness of the disease in question, and mixing patterns in the population. The effects of vaccination, for example, which alters the susceptibility status of the population and, potentially, the strength of infectiousness and other factors, can thereby be incorporated accounting for both the direct effects of vaccination (i.e. protection conferred upon the vaccinated individual), and its indirect effects (i.e. herd immunity). However, deterministic compartmental transmission models, which are in essence cohort models, require a number of simplifying assumptions which can be unacceptable in specific situations [68,71]. The simplest models assume complete uniformity in the population, instantaneous contacts and infection, instantaneous and “memoryless” mixing, uniformity in infection events, and uniformity in recovery from infection [71]. These assumptions often will not be supportable when providing information to decision makers about the economic consequences of related interventions. Some of the assumptions in transmission models can be relaxed with more sophisticated transmission modelling techniques. Techniques specifically designed to address those issues provide a more convenient and accurate modelling framework [71].

## 11 Handling variability and uncertainty

### 11.1 Types of variability

There are two major kinds of variability in economic models [84]. There is a variability that *per se* cannot be avoided and that results from differences in the values of variables for different patients. A second form of variability results from the existence of “incomplete information” in the model, i.e. from the assumptions on which the model is based. The first form of variability depends on the characteristics of the intervention, the patients, the context and the unit costs. Note that these elements may be known with certainty (e.g. age, gender) or there may be uncertainty about them (e.g. the *per diem* cost of a hospital stay, effect estimate from a clinical study). If the values of variables are known, it is better to talk about heterogeneity than variability. The second form of variability reflects the fact that models can not completely capture and reflect reality due to its complexity. Therefore, models have to rely on simplifying assumptions in order to remain manageable, traceable and transparent. This kind of variability is inherent in modelling. Both forms of variability have to be quantified and the modelling must take account of them.

### 11.2 Quantifying uncertainty

Generally speaking, the models are supplied with data from different sources, e.g. clinical trials, observational studies, data base analyses, surveys, etc. Thus, there is some degree of uncertainty in all models [102]. The sources of this uncertainty can be categorized in a number of ways [37,50,103-105].

(1) The uncertainty of model parameters may result from their variability and the fact that they are estimated from samples with finite sample size. Thus, parameter uncertainty reflects our inability to determine these values precisely, such as those pertaining to the efficacy of the intervention or to the relevant resource utilization [106]. This type of uncertainty is often described by means of confidence intervals or other statistical approaches to quantify variance. These measures reflect the observed uncertainty in that parameter within the source data.

(2) Costs may vary because there is uncertainty in some of their determinants such as the dose consumed over time. Also, the model may be stochastic (i.e. uses random numbers in Monte Carlo simulations). To minimize this type of variability, various techniques are deployed to eliminate as much of it as possible. These variance reduction techniques (e.g. cloning) are dealt with extensively in the literature [73,107,108]. One common technique is to increase the sample size by re-sampling data from given patients (bootstrapping) or by running the model several times. The remaining variation after all the variance reduction techniques have been deployed must be quantified.

(3) The heterogeneity of the patient population, such as age, gender, relevant clinical history has an important impact on the applicability of the economic evaluation's results. Although the characteristics of the patients may vary across the population, this does not *per se* imply uncertainty [103]. There can be uncertainty, however, because of insufficient information on these characteristics, and if so this should be quantified along with other parameter uncertainties.

(4) Equally the variability in model assumptions, which was described in the previous chapter, causes uncertainty and has to be taken into account. The modeller has to disclose and document the decisions regarding the choice of a modelling structure and the underlying assumptions. For instance, these may be related to the way the healthcare intervention is conceptualized, the clinical events included, the functional relations, the choice of costs and benefits to be addressed, as well as the methods of measurement and assessment of outcomes [45,102,103,109].

(5) Fixed factors such as the discounting rate or the selected time horizon can also fluctuate within a range of values and thus vary.

### **11.3 Handling uncertainty and its effects on model results**

Parameter and other uncertainty that is irreducible have to be quantified. All analyses undertaken to do so should be fully documented in terms of the ranges for the parameter values used and assumptions made.

Uncertainties in the model inputs have to be considered and their impact on the model results quantified using the proper methods [26,103,110]. The three main methods used in health economic evaluations to test robustness of model results in the light of uncertainty are univariate deterministic, multivariate deterministic, and multivariate probabilistic sensitivity analysis. Other methods exist [107,111] but are rarely applied in health technology assessments. Multivariate analyses are most useful if the analysis is carried out for specific scenarios defined by a set of values for these elements.

Univariate sensitivity analysis addresses the range in outcomes resulting from varying one of the model parameters across a plausible range of uncertainty, while all other parameters are kept at their original assigned values (it should be noted that this may not take into account associations between inputs). This can give an idea of how large the impact of that uncertainty is, but obviously cannot fully quantify the uncertainty in the results. Univariate sensitivity analyses are extremely helpful if a decision maker has assumptions regarding one important parameter that differ from the base case analysis.

Multivariate sensitivity analysis attempts to make the assessment of uncertainty more complete by changing more than one parameter at a time and investigating their joint impact. Although this can be very useful for examining scenarios when the uncertainty is of interest, it

is time consuming when each parameter has a broad range and not very accurate if associations between parameter values are not taken into account.

In probabilistic sensitivity analysis, each parameter that is subject to uncertainty is assigned a distribution and these distributions are simultaneously sampled to create a set of inputs for one run of the model [45]. This is replicated many times and the resulting set of results provides an idea of the total impact of parameter uncertainty. For this to reasonably reflect uncertainty, it is important to choose distributions that accurately represent the uncertainty around the parameter estimates, and to correctly reflect how these correlate with each other. Otherwise, parameter combinations that rarely (or never) occur in reality may be over-represented.

It is recommended that univariate sensitivity analyses are not replaced but complemented by multivariate and probabilistic sensitivity analyses. For univariate and multivariate analyses, the results should be reported in both tabular form and as a tornado diagram giving the range of results for the input range(s). In case of probabilistic sensitivity analyses, results should be presented as the cumulative distribution of results. Cost-effectiveness acceptability curves [112] are not requested in IQWiG evaluations, given that these analyses are not about threshold cost-effectiveness ratios. Structural sensitivity analyses should be undertaken to explore the impact of varying structural assumptions.

Models that go beyond the duration of underpinning RCTs are subject to greater uncertainty than models that are limited to the duration of an RCT. Accordingly, it is of increasing importance to adequately test and document the robustness of model results with longer modelling time spans.

## 12 Validation

The approaches to validation have been classified in many ways, which will not be explored in detail here. However, in the context of modelling, clarification is needed as to which steps validation must consist of.

One key element is to examine whether the model makes sense. This so-called face validity relates to the influence diagram, the model concept, the data acquisition, the processing of functional relations and the choice of modelling technique. Each step must be carried out with prior knowledge, best available information and taking account of best practice. A second key element in validation is whether the model is implemented correctly (also called technical validation). This aspect also covers whether the model is operating in the manner intended – that the logic of the program is correctly implemented, with a minimum of programming errors. The third key aspect is external validation of the model – establishing whether the model correctly reflects the “real world”, within the defined limitations of the model. This is extremely important since a technically good, bug-free model with strong face validity is of little value for decision making if it does not accurately reflect the system and its behaviour.

A simulation model that is valid for one purpose may not be valid for another [113]. The external validation process must therefore cover all intended uses of the model, and if it is put to different purposes later, then renewed validation should take place. There is disagreement over the procedure for validating a model, but there are some basic steps that have to be taken [114].

The design and structure of the model should be reviewed by other experts in the field. This should be done for the influence diagram and model concept. The influence diagram is an important aspect of face validity. As it displays the modeller’s understanding of the illness and its management, it provides a transparent tool for clinical and other experts to assess the planned model.

All equations to be used in the model should be tested separately before being entered into the model. This is to ensure that they are correctly expressed and give the right results for entered values. The equations should also be calculated with extreme values on the borders of the model scope to make sure that there are no problems in those regions.

It is a huge task to test all aspects of the model when it is completed, due to the complexities and interactions. The only sensible approach is to keep testing the model as it evolves. This is best done by constructing the model in modular fashion. Each module is tested when it is developed. The module testing consists of checking that its output is correct, given the inputs to the module (e.g. entities and data being entered in a discrete event simulation). Any changes to a module already tested trigger new testing of that module. As the modules are added on, continuous testing should take place, thus ensuring consistent behaviour. In an individual-level simulation (Markov modelling with first order Monte Carlo simulation or discrete event simulation), a basic method is to test single, well-defined patients to make sure

the paths and events they experience are the expected ones. Moreover, the same approach chosen to test the single modules should be used for the whole model.

A good way of externally validating a model is to run it using the inputs from historical data, such as completed clinical trials, registries or outcomes studies, from which data have not been used in the tested model. The model results are then compared with the known ones [9]. By testing trials that have not been used as sources for inputs, the ability of the model to reproduce reality is verified. This verification must be tempered, however, as trials themselves are substantial deviations from reality. Ideally, the same process should take place with data from actual practice, although these are rarely available.

## 13 Model documentation

A detailed technical report describing all the modelling steps from development of the initial influence diagram to final validation is required. In addition, a fully executable version of the model must be made available, along with a user manual. In line with other suggested guidelines [6,12,24,26,110], documentation on the model should include the following:

- The influence diagram used to guide model development
- Details of the model concept
  - Description of the population(s) considered in the evaluation, including subgroups
  - Description of the evaluated health technologies
  - Selection of model settings (simulation size, time horizon, discounting rates, etc.) and justification
  - Review of economic evaluations in the therapeutic area in question
- Description of all data sources. Justification for selection of data sources must be provided.
- Details of all functional relationships used in the model. If they were custom- developed for the model, the methods used must be provided in detail.
- Listing of all assumptions related to both data source and model structure. Especially important is a detailed account of any assumptions and techniques used to project beyond the source data.
- Rationale for the modelling technique adopted
  - Description of how the technique conforms with the mandatory features
- Overview of the validation techniques used and results
- Detailed results, including an assessment of the impact of:
  - Use in relevant population subgroups
  - Uncertainty in input data
  - If probabilistic sensitivity analyses are conducted, include specification of
    - the probability distributions used and sources
    - correlations among input parameters

- any structural variants
- Interpretation of results, including a description of limitations of the approach adopted.

An electronic version of the model must be provided on the understanding that the model will enter the public domain, and may be adapted for use in future evaluations. The electronic model must be fully accessible and allow reviewers and the public to view all formulas and relationships included in the analysis and to run the model using different input data. To facilitate review of the model, the electronic model should be accompanied by a user's guide describing the software and hardware required to use the model, how model inputs can be modified, where these inputs can be found in the model, and how the model may be run and results extracted.

## 14 References

1. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Econ* 2006; 15(7): 677-687.
2. Lang DL, Lopert R, Hill SR. Use of pharmacoeconomics in prescribing research; part 5: modelling; beyond clinical trials. *J Clin Pharm Ther* 2003; 28(5): 433-439.
3. Siebert U, Sroczynski G. Effectiveness and cost-effectiveness of initial combination therapy with interferon/peginterferon plus ribavirin in patients with chronic hepatitis C in Germany: a health technology assessment commissioned by the German Federal Ministry of Health and Social Security. *Int J Technol Assess Health Care* 2005; 21(1): 55-65.
4. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004; 9(2): 110-118.
5. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T et al. Modelling in economic evaluation: an unavoidable fact of life. *Health Econ* 1997; 6(3): 217-227.
6. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C et al. Principles of good practice for decision analytic modeling in health care evaluation: report of the ISPOR task force on good research practices-modeling studies. *Value Health* 2003; 6(1): 9-17.
7. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; 15(12): 1295-1310.
8. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? *Eur J Health Econ* 2003; 4(3): 143-150.
9. Eddy DM. Accuracy versus transparency in pharmacoeconomic modelling. *Pharmacoeconomics* 2006; 24(9): 837-844.
10. Ades AE, Claxton K, Sculpher MJ. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. *Health Econ* 2006; 15(4): 373-381.
11. Chapman GB, Sonnenberg FA. Decision making in health care: theory, psychology, and applications. Cambridge: Cambridge University Press; 2000.
12. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.

13. Philips Z, Bojke L, Sculpher MJ, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006; 24(4): 355-371.
14. Eddy DM. *Technology assessment: the role of mathematical modeling*. Washington: National Academy Press; 1985.
15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Allgemeine Methoden zur Bewertung von Verhältnissen zwischen Nutzen und Kosten: Version 1.0*. Köln: IQWiG; 2009.
16. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Arbeitspapier Kostenbestimmung: Version 1.0*. Köln: IQWiG; 2009.
17. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2000.
18. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Methoden: Version 3.0* [online]. 2008 May 27 [cited 2009 Oct 09]. URL: [http://www.iqwig.de/download/IQWiG\\_Methoden\\_Version\\_3\\_0.pdf](http://www.iqwig.de/download/IQWiG_Methoden_Version_3_0.pdf).
19. Leidl R, Graf von der Schulenburg JM, Wasem J (Ed). *Ansätze und Methoden der ökonomischen Evaluation: eine internationale Perspektive*. Baden-Baden: Nomos; 1999. (Health Technology Assessment).
20. Nuijten MJC. The selection of data sources for use in modelling studies. *Pharmacoeconomics* 1998; 13(3): 305-316.
21. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application. *Pharmacoeconomics* 2000; 17(5): 461-477.
22. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Economic evaluation using patient level data. In: *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2005. p. 247-255.
23. Ramsey SD, Willke R, Briggs A, Brown R, Buxton M, Chawla A et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health* 2005; 8(5): 521-533.
24. Soto J. Health economic evaluations using decision analytic modeling. *Int J Technol Assess Health Care* 2002; 18(1): 94-111.
25. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-149.

- 
26. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: NICE; 2004. URL: [http://www.nice.org.uk/niceMedia/pdf/TAP\\_Methods.pdf](http://www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf).
  27. Song F, Altman DG, Glenny A, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published metaanalyses. *BMJ* 2003; 326(7387): 472.
  28. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006; 101(474): 447-459.
  29. Lu G, Ades AE, Sutton AJ, Cooper NJ, Briggs AH, Caldwell DM. Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Stat Med* 2007; 26(20): 3681-3699.
  30. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.
  31. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002; 21(16): 2313-2324.
  32. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.
  33. Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med* 2007; 26(6): 1237-1254.
  34. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008; 26(9): 753-767.
  35. Sculpher MJ, Drummond MF. Analysis sans frontières: can we ever make economic evaluations generalisable across jurisdictions? *Pharmacoeconomics* 2006; 24(11): 1087-1099.
  36. Barbieri M, Drummond M, Willke R, Chancellor J, Jolain B, Towse A. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalizability. *Value Health* 2005; 8(1): 10-22.
  37. Drummond MF, McGuire AE. *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001.
  38. Agency for Healthcare Research and Quality. *Registries for evaluating patient outcomes: a user's guide*; AHRQ publication no. 07-EHC001-1. Rockville: AHRQ; 2007. URL: <http://effectivehealthcare.ahrq.gov/repFiles/PatOutcomes.pdf>.
  39. Statistisches Bundesamt. Statistisches Bundesamt Deutschland [online]. [cited 2009 Oct 09]. URL: <http://www.destatis.de>.

- 
40. The Human Mortality Database [online]. [cited 2009 Oct 09]. URL: <http://www.mortality.org>.
41. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 1996; 276(15): 1253-1258.
42. Drummond MF, Sculpher MJ. Common methodological flaws in economic evaluations. *Med Care* 2005; 43(7 Suppl): 5-14.
43. Commonwealth Department of Health and Ageing. Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee: including major submissions involving economic analyses. Canberra: Commonwealth of Australia; 2002. URL: [http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-general-pubs-guidelines-index.htm/\\$FILE/guidelines.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-general-pubs-guidelines-index.htm/$FILE/guidelines.pdf).
44. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Collection and analysis of data. In: *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2000. p. 234-235.
45. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. Oxford: Oxford University Press; 1996.
46. Ramsey SD, McIntosh M, Etzioni R, Urban N. Simulation modeling of outcomes and cost effectiveness. *Hematol Oncol Clin North Am* 2000; 14(4): 925-938.
47. Nixon RM, Thompson SG. Methods for incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. *Health Econ* 2005; 14(12): 1217-1229.
48. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002; 11(5): 415-430.
49. Vazquez-Polo FJ, Negrin Hernandez MA, Lopez-Valcarcel BG. Using covariates to reduce uncertainty in the economic evaluation of clinical trial data. *Health Econ* 2005; 14(6): 545-557.
50. Briggs AH, Claxton K, Sculpher M. *Decision modelling for health economic evaluation: handbooks for health economic evaluation*. Oxford: Oxford University Press; 2006.
51. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer; 2001.

- 
52. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004; 13(5): 461-475.
53. Kleinbaum DG, Klein M. *Survival analysis: a self-learning text*. New York: Springer; 1996.
54. Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. Hoboken: Wiley; 2002.
55. Lawless JF. *Statistical models and methods for lifetime data*. Hoboken: Wiley; 2003.
56. Nixon RM, Thompson SG. Parametric modelling of cost data in medical studies. *Stat Med* 2004; 23(8): 1311-1331.
57. Kilian R, Matschinger H, Löffler W, Roick C, Angermeyer MC. A comparison of methods to handle skew distributed cost variables in the analysis of the resource consumption in schizophrenia treatment. *J Ment Health Policy Econ* 2002; 5(1): 21-31.
58. Manca A, Rice N, Sculpher MJ, Briggs AH. Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. *Health Econ* 2005; 14(5): 471-485.
59. Mullahy J. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *J Health Econ* 1998; 17(3): 247-281.
60. Cook JR, Drummond M, Glick H, Heyse JF. Assessing the appropriateness of combining economic data from multinational clinical trials. *Stat Med* 2003; 22(12): 1955-1976.
61. Robins JM, Blevins D, Ritter G, Wulfsohn M. G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients. *Epidemiology* 1992; 3(4): 319-336.
62. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran ME, Berry D (Ed). *Statistical models in epidemiology: the environment and clinical trials*. New York: Springer; 1999. p. 95-134. (The IMA volumes in mathematics and its applications; volume 116).
63. Robins JM, Hernan MA, Siebert U. Estimations of the effects of multiple interventions. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL (Ed). *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*. Genf: WHO; 2004. p. 2191-2230.
64. Draper NR, Smith H. *Applied regression analysis*. New York: Wiley; 1998.

- 
65. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999; 10(1): 37-48.
66. Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci USA* 2002; 99(Suppl 3): 7280-7287.
67. Cooper K, Brailsford SC, Davies R. Choice of modeling technique for evaluating health care interventions. *J Oper Res Soc* 2007; 58(2): 168-176.
68. Koopman JS, Jacquez G, Chick SE. New data and tools for integrating discrete and continuous population modeling strategies. *Ann N Y Acad Sci* 2001; 954: 268-294.
69. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. *Pharmacoeconomics* 2008; 26(2): 131-148.
70. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. *Pharmacoeconomics* 2006; 24(11): 1043-1053.
71. Koopman J. Modeling infection transmission. *Annu Rev Public Health* 2004; 25: 303-326.
72. Caro JJ. Pharmacoeconomic analyses using discrete event simulation. *Pharmacoeconomics* 2005; 23(4): 323-332.
73. Shechter SM, Schaefer AJ, Braithwaite RS, Roberts MS. Increasing the efficiency of Monte Carlo cohort simulations with variance reduction techniques. *Med Decis Making* 2006; 26(5): 550-553.
74. Graf von der Schulenburg JM, Greiner W, Jost F, Klusen N, Kubin M, Leidl R et al. Deutsche Empfehlungen zur gesundheitsökonomischen Evaluation: dritte und aktualisierte Fassung des Hannoveraner Konsens. *Gesundheitsökonomie & Qualitätsmanagement* 2007; 12(5): 285-290.
75. Kuntz KM, Weinstein MC. Modelling in economic evaluation. In: Drummond MF, McGuire AE (Ed). *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001. p. 141-171.
76. Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required? *Value Health* 2006; 9(4): 244-252.
77. Hollingworth W, Spackman DE. Emerging methods in economic modeling of imaging costs and outcomes a short report on discrete event simulation. *Acad Radiol* 2007; 14(4): 406-410.
78. Lee RC, Donaldson C, Cook LS. The need for evolution in healthcare decision modeling. *Med Care* 2003; 41(9): 1024-1033.

- 
79. Banks J, Carsons JS, Nelson BL, Nicol DM. Discrete-event system simulation. New Jersey: Pearson Prentice Hall; 2005.
80. Macal CM, North MJ. Tutorial on agent-based modeling and simulation [online]. In: Kuhl ME, Steiger NM, Armstrong FB, Joines JA (Ed). Proceedings of the 2005 Winter Simulation Conference; 2005 Dec 04-07; Orlando, USA. 2005. p. 2-15. [cited 2009 Oct 09]. URL: <http://www.informs-sim.org/wsc05papers/002.pdf>.
81. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making* 1983; 3(4): 419-458.
82. Briggs AH, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998; 13(4): 397-409.
83. Siebert U, Mühlberger N, Schöffski O. Evidenzsynthese: Meta-Analysen und Entscheidungsanalysen. In: Graf von der Schulenburg JM (Ed). *Gesundheitsökonomische Evaluationen*. Berlin: Springer; 2007. p. 261-310.
84. Hunink MG, Glasziou PP, Siegel JE, Weeks JC, Pliskin JS, Elstein AS et al. *Decision making in health and medicine: integrating evidence and values*. Cambridge: Cambridge University Press; 2001.
85. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ* 2003; 12(10): 837-848.
86. Davies HTO, Davies R. Simulating health systems: modelling problems and software solutions. *Eur J Oper Res* 1995; 87(1): 35-44.
87. Fishman GS. *Discrete-event simulation: modeling, programming, and analysis*. Berlin: Springer; 2001.
88. Bratley P, Fox BL, Schrage LE. *A guide to simulation*. New York: Springer; 1987.
89. Kelton WD, Sadowski RP, Sturrock DT. *Simulation with ARENA*. Boston: McGraw-Hill; 2007.
90. Alexopoulos C, Goldsman D, Fontanesi J, Sawyer M, De Guire M, Kopald D et al. A discrete-event simulation application for clinics serving the poor [online]. In: Peters BA, Smith JS, Medeiros DJ, Rohrer MW (Ed). Proceedings of the 2001 Winter Simulation Conference; 2001 Sep 09-12; Arlington, USA. 2001. p. 1386-1391. [cited 2009 Oct 09]. URL: <http://www.informs-sim.org/wsc01papers/188.PDF>.
91. Cahill W, Render M. Dynamic simulation modeling of ICU bed availability [online]. In: Farrington PA, Nembhard HB, Sturrock GW, Evans GW (Ed). Proceedings of the 1999

- Winter Simulation Conference; 1999 Dec 05-08; Squaw Peak, USA. 1999. p. 1573-1576. [cited 2009 Oct 09]. URL: <http://www.informs-sim.org/wsc99papers/228.PDF>.
92. Gray J, Geva A, Zheng Z, Zupancic JA. CoolSim: using industrial modeling techniques to examine the impact of selective head cooling in a model of perinatal regionalization. *Pediatrics* 2008; 121(1): 28-36.
93. Jun JB, Jacobson SH, Swisher JR. Applications of discrete-event simulation in health care clinics: a survey. *J Oper Res Soc* 1999; 50(2): 109-123.
94. Proctor T. Simulation in hospitals. *Health Manpow Manage* 1996; 22(5): 40-44.
95. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993; 13(4): 322-338.
96. Duffie D, Glynn P. Estimation of continuous-time Markov processes sampled at random time intervals. *Econometrica* 2004; 72(6): 1773-1808.
97. Stewart WJ. Introduction to the numerical solution of Markov chains. Princeton: Princeton University Press; 1994.
98. Bilge U, Saka O. Agent-based simulations in health care. In: Hasman A, Haux R, Van der Lei J, De Clercq E, Roger France FH (Ed). *Ubiquity: technologies for better health in aging societies; proceedings of MIE2006*. Amsterdam: IOS Press; 2006. p. 699-704. (Studies in health technology and informatics; volume 124).
99. Burke DS, Epstein JM, Cummings DAT, Parker JI, Cline KC, Singa RM et al. Individual-based computational modeling of smallpox epidemic control strategies. *Acad Emerg Med* 2006; 13(11): 1142-1149.
100. Gorman DM, Mezic J, Mezic I, Gruenewald PJ. Agent-based modeling of drinking behavior: a preliminary model and potential applications to theory and practice. *Am J Public Health* 2006; 96(11): 2055-2060.
101. Anderson RM, May RM. *Infectious disease of humans: dynamics and control*. Oxford: Oxford University Press; 1991.
102. Bojke L, Claxton K, Palmer S, Sculpher MJ. Defining and characterising structural uncertainty in decision analytic models [online]. 2006 Mar [cited 2009 Oct 09]. (CHE Research Papers; volume 9). URL: <http://www.york.ac.uk/inst/che/pdf/rp9.pdf>.
103. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000; 17(5): 479-500.
104. Cantor SB, Ganiats TG. Incremental cost-effectiveness analysis: the optimal strategy depends on the strategy set. *J Clin Epidemiol* 1999; 52(6): 517-522.

105. Karlsson G, Johannesson M. Cost-effectiveness analysis and capital costs. *Soc Sci Med* 1998; 46(9): 1183-1191.
106. O'Hagan A, McCabe C, Akehurst R, Brennan A, Briggs A, Claxton K et al. Incorporation of uncertainty in health economic modelling studies. *Pharmacoeconomics* 2005; 23(6): 529-536.
107. Law AM, Kelton WD. *Simulation modelling and analysis*. New York: McGraw Hill; 2000.
108. Ross SM. *Simulation*. Oxford: Elsevier Academic Press; 2006.
109. Drummond MF, Brandt A, Luce B, Rovira J. Standardizing methodologies for economic evaluation in health care: practice, problems, and potential. *Int J Technol Assess Health Care* 1993; 9(1): 26-36.
110. Collège des Économistes de la Santé. French guidelines for the economic evaluation of health care technologies [online]. 2004 Sept [cited 2009 Oct 09]. URL: [http://www.ces-asso.org/docs/France\\_Guidelines\\_HE\\_Evaluation.PDF](http://www.ces-asso.org/docs/France_Guidelines_HE_Evaluation.PDF).
111. Oberkampf WL, DeLand SM, Rutherford BM, Diegetd KV, Alvine KF. Error and uncertainty in modeling and simulation. *Reliability Engineering and System Safety* 2002; 75(3): 333-357.
112. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves-facts, fallacies and frequently asked questions. *Health Econ* 2004; 13(5): 405-415.
113. Law AM, McComas MG. How to build valid and credible simulation models [online]. In: Peters BA, Smith JS, Medeiros DJ, Rohrer MW (Ed). *Proceedings of the 2001 Winter Simulation Conference*; 2001 Sep 09-12; Arlington, USA. 2001. p. 22-29. [cited 2009 Oct 09]. URL: <http://www.informs-sim.org/wsc01papers/004.PDF>.
114. Sargent RG. Validation and verification of simulation models [online]. In: Ingalls RG, Rossetti MD, Smith JS, Peters BA (Ed). *Proceedings of the 2004 Winter Simulation Conference*; 2004 Dec 05-08; Washington, USA. 2004. p. 17-28. [cited 2009 Oct 09]. URL: <http://www.informs-sim.org/wsc04papers/004.pdf>.