

Technischer Anhang

Modellierung

Version 1.0

09.10.2008

Kontakt:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Dillenburger Straße 27

D-51105 Köln

Tel.: +49 (0)221 / 35685-0

Fax: +49 (0)221 / 35685-1

E-mail: knb-methoden@iqwig.de

Inhaltsverzeichnis

Inhaltsverzeichnis.....	i
1 Einleitung	1
2 Definition	2
3 Zweck	3
4 Vorgehen	5
5 Einflussdiagramme.....	9
5.1 Definition	9
5.2 Zweck.....	9
5.3 Komponenten	9
6 Modellkonzept.....	11
6.1 Definition	11
6.2 Zweck.....	11
6.3 Komponenten	11
6.4 Validierung.....	13
7 Erforderliche Modelleigenschaften.....	14
7.1 Hinreichende Detailliertheit	14
7.2 Perspektive	15
7.3 Zeithorizont	15
7.4 Behandlungsdauer	16
7.5 Diskontierung	16
8 Datenbewertung	17
8.1 Ziele.....	18
8.2 Potenzielle Datenquellen.....	18
8.2.1 Klinische Studien	18
8.2.2 Epidemiologische Studien.....	20
8.2.3 Prozessdatenbanken	21
8.2.4 Register.....	21
8.2.5 Kompilierte Statistiken.....	22
8.2.6 Expertenmeinung	22
9 Funktionale Beziehungen.....	24
9.1 Definition	24
9.2 Zweck.....	24
9.3 Allgemeine Konzepte.....	24
10 Modellierungstechniken.....	29
10.1 Aspekte der Wahl der Modellierungstechnik.....	30
10.1.1 Kohortenmodelle versus Individualmodelle	30
10.1.2 Ereignisbasierte versus zustandsübergangsbasierte Simulation.....	31
10.2 Techniken	34
10.2.1 Entscheidungsbaumanalysen.....	34
10.2.2 Diskrete Ereignissimulation	35
10.2.3 Markov-Modelle.....	39
10.2.4 Agentenbasierte Simulation	43
10.2.5 Transmissionsmodelle.....	44
10.3 Empfohlene Technik	45

11	Umgang mit Unsicherheit und Variabilität	47
11.1	Arten.....	47
11.2	Quantifizierung von Unsicherheit	48
11.3	Empfehlung	50
12	Validierung.....	51
13	Modelldokumentation	54
	Literaturverzeichnis.....	57

1 Einleitung

Abschätzungen der ökonomischen Nettoeffekte von Gesundheitstechnologien, einschließlich deren möglicher Budgetauswirkungen, sind unentbehrlich. In den meisten Fällen werden ökonomische Daten, die in klinischen Studien gesammelt wurden, für sich genommen nicht für eine vollständig fundierte Entscheidungsfindung ausreichend sein. Kostendaten aus klinischen Studien sind meistens nicht besonders brauchbar, da die Studienbedingungen von denen in der realen Praxis abweichen, wodurch die Übertragbarkeit leidet. Darüber hinaus sind sie häufig weder spezifisch für Deutschland noch adäquat auf Deutschland übertragbar und liefern zumeist keine Information hinsichtlich der langfristigen Konsequenzen der Einführung einer neuen Technologie [1,2]. Tatsächlich könnten ökonomische Daten aus klinischen Studien auch überhaupt nicht verfügbar sein. Deswegen ist die Modellierung ökonomischer Outcomes (Zielgrößen) eine wesentliche Komponente der Evaluation.

Dieser technische Anhang beschreibt Modellierungsansätze im Gesundheitswesen und gibt Handlungsempfehlungen für die Entwicklung von ökonomischen Modellen, die Modellierungstechnik, die Modellvalidierung, die Unsicherheitsanalyse (engl. „uncertainty analysis“) und den Bericht von Modellierungsstudien.

Obwohl das IQWiG prinzipiell aufgefordert werden kann, verschiedene Arten von Gesundheitstechnologien (z. B. präventive, diagnostische und therapeutische Verfahren) zu bewerten, wird in diesem technischen Anhang schwerpunktmäßig auf technische Aspekte bei der ökonomischen Modellierung von therapeutischen Verfahren Bezug genommen.

2 Definition

Für den Begriff „Modell“ im Kontext des Gesundheitswesens finden sich verschiedene Definitionen. Modelle sind analytische Werkzeuge, die dazu verwendet werden, Systeme in der realen Welt zu verstehen, verschiedene Outcomes bei einem gegebenen Satz von Inputparametern zu schätzen und die Auswirkungen von Veränderungen auf das System zu modellieren. Daraus ergibt sich, dass jede über die direkte Anwendung beobachteter Daten hinaus gehende Evaluation als Modell betrachtet werden kann [3], und sogar die direkte Anwendung beinhaltet normalerweise in einem gewissen Ausmaß statistische Modellierung. Selbstverständlich können Modelle die Wirklichkeit nicht perfekt abbilden: Sie basieren auf einem reduzierten Satz von Komponenten und erfordern vereinfachende Annahmen [4]. Dennoch ist die Validität eines Modells entscheidend, was bedeutet, dass es das von ihm repräsentierte System hinreichend widerspiegeln muss.

Die ISPOR (International Society for Pharmacoeconomics and Outcomes Research) definiert „Modellierung“ als „an analytic methodology that accounts for events over time and across populations, that is based on data drawn from primary and/or secondary sources“ [5]. Eine aktuelle Beschreibung von Modellierungstechniken gibt eine ähnliche Definition: „a formal quantified comparison of health technologies synthesizing sources of evidence on costs and benefits“ [6]. Beide Definitionen sind vereinbar mit den Anforderungen an ökonomische Evaluationen, die vom IQWiG durchgeführt werden.

3 Zweck

Es wurden verschiedene Situationen beschrieben, in denen Modellierung bei ökonomischen Evaluationen im Gesundheitswesen Anwendung finden sollte [7]. Der ultimative Zweck der Modellierung von ökonomischen Effekten einer Gesundheitstechnologie ist, eine solide Basis für Entscheidungen hinsichtlich der Angemessenheit der Erstattung einer Technologie zu einem bestimmten Preis bereitzustellen. Im Kontext der Entscheidungsfindung in Deutschland basieren die Gesundheitseffekte einer Intervention auf den patientenrelevanten Nutzeneffekten, die vom IQWiG unter Verwendung der internationalen Standards der evidenzbasierten Medizin analysiert werden. Patientenrelevante Nutzeneffekte diagnostischer und therapeutischer Interventionen werden in kontrollierten klinischen Studien ermittelt. Valide Modelle können ggf. Vorhersagen darüber treffen, wie sich die in klinischen Studien ermittelten Nutzeneffekte unter anderen Bedingungen oder Zeitperspektiven darstellen würden. Modelle können jedoch in keinem Fall neue Nutzeneffekte generieren. Hierdurch konzentriert sich das Ziel der Modellierung in Deutschland hauptsächlich auf die ökonomischen Konsequenzen der Gesundheitstechnologie unter verschiedenen Szenarien. Wenn die medizinische Nutzenbewertung, welche vor der ökonomischen Bewertung erfolgt, bereits eine Modellierung erforderlich macht, werden Modellierungstechniken sowohl zur Schätzung der gesundheitlichen Konsequenzen als auch der Kosten eingesetzt. Dies könnte der Fall sein, wenn sich die verglichenen Gesundheitstechnologien auf mehrere Aspekte der Gesundheit gleichzeitig auswirken, welche dann in einem gemeinsamen Score kombiniert werden müssen (z. B. qualitätsadjustierte Lebensjahre oder andere zusammenfassenden Maße) (siehe Kapitel 2.3.1 im Hauptdokument), oder der medizinische Nutzen in eine Kardinalskala transformiert werden muss, die den Wert (engl. „value“) dieses Nutzens widerspiegelt. Dieser Transfer kann Modellierung mit einbeziehen, sofern beispielsweise ein längerer Zeitraum betrachtet werden soll (siehe Kapitel 2.3.1 im Hauptdokument).

Diese Information wird dazu verwendet, die Technologie in Bezug auf die Effizienzgrenze (siehe Kapitel 3) zu positionieren und die Budgetauswirkungen zu schätzen. Dies bedeutet nicht, dass Gesundheitsoutcomes nicht in ökonomische Modelle des IQWiG integriert werden, sondern vielmehr, dass Modellierung nicht vornehmlich zur Abschätzung der Gesundheitsoutcomes per se verwendet wird. Gesundheitsoutcomes müssen vielmehr Teil der Modellierung sein, da ohne diese Komponente keine aussagekräftige und reliable

Abschätzung der Kosten möglich ist und die ökonomische Evaluation auf den unzureichenden Vergleich der Akquisitionskosten konkurrierender Therapieoptionen reduziert wird.

Kritiker des Einsatzes von Modellierung in Evaluationen haben argumentiert, dass diese analytischen Strukturen hinsichtlich der Wahl der Eingangsparameter, des strukturellen Designs und der getroffenen Annahmen verzerrungsanfällig und möglicherweise unzugänglich sind – eine sogenannte *black box*, die nur die Modellentwickler selbst verstehen. Ein gegebenes Modell kann all diese negativen Eigenschaften aufweisen, wobei es aber möglich ist, ein klar verständliches Modellkonzept mit expliziten Annahmen und gemäß Datenlage bestmöglichen Inputparametern zu entwerfen [8]. Modellierung ermöglicht ebenfalls die Exploration von verschiedenen Einsatzszenarien und -strategien [4]. Damit die Modellierungsstudien, die vom IQWiG durchgeführt werden, von Nutzen sind, ist jedoch Folgendes entscheidend:

1. vollständige Transparenz mit klar beschriebenen und begründeten Modellinputs und Annahmen,
2. hinreichende Tiefe zur ausreichenden Abbildung der modellierten Erkrankung, der mit ihr assoziierten Kosten und der betreffenden Versorgungsmaßnahmen,
3. ausreichende Flexibilität zur Berechnung multipler Szenarien bei variierenden Annahmekonstellationen und Settings,
4. Möglichkeit der Bestimmung der Unsicherheit in den vorhergesagten Kosten und
5. Verwendung von Daten, die für Deutschland relevant sind. Dabei sollten nicht nur Kosten, sondern auch praxisübliche Behandlungsmuster, Demografie und Epidemiologie berücksichtigt werden.

Modelle müssen außerdem einer strengen Validierung unterzogen werden, sowohl im Sinne der Integrität der internen Berechnungen als auch im Hinblick auf die externe Validität. Deshalb sollten Modelle des IQWiG einem Review unterzogen werden, bei dem Zugang sowohl zu den relevanten technischen Dokumenten als auch zu einer voll funktionsfähigen und evaluierbaren elektronischen Version des Modells gewährt wird.

4 Vorgehen

Eine Beschreibung des Vorgehens bei der Modellentwicklung, Validierung, Analyse und Berichterstattung findet sich im anschließenden Teil dieses technischen Anhangs. Im Folgenden sind kurz die wesentlichen Schritte dargestellt, die für die Modellentwicklung empfohlen werden:

1. Definition der *Forschungsfragestellung(en)*, deren Beantwortung durch das Modell erwartet wird. Dies beinhaltet:
 - a. Welche therapeutischen Handlungsalternativen werden bewertet?
 - b. Welche Population oder Populationen sind relevant für die Analyse?
 - c. Welches sind die relevanten Kostenkategorien, die mit der betreffenden Indikation und den evaluierten Interventionen in Zusammenhang stehen? (Siehe auch Kapitel 4)?
2. Entwurf eines *Einflussdiagramms* (engl. „*influence diagram*“), welches grafisch den Zusammenhang zwischen den Schlüsselfaktoren der Fragestellung sowie deren Interaktion beschreibt. Das Einflussdiagramm dient dem Zweck, einen Überblick über die Erkrankung sowie deren ökonomische Aspekte zu erhalten und die entscheidenden Komponenten der Bezüge im geplanten Modell zusammenzufassen.
3. Eng im Zusammenhang mit der Entwicklung des Einflussdiagramms steht die Spezifizierung des Modellkonzepts in Form von Flussdiagrammen, welche als Konstruktionsplan für das Modell dienen. Diese Flussdiagramme sollten erläutern, wie die Gesundheitseffekte der evaluierten Handlungsalternativen mit den konkreten Gesundheitsoutcomes und dem Ressourcenverbrauch zusammenhängen, sowie deren Projektion über den Zeithorizont des Modells verdeutlichen. Das Modellkonzept sollte erläutern, wie mit den wechselseitigen Beziehungen im Einflussdiagramm umgegangen wird, und Begründungen für das gewählte Vorgehen liefern. Das Konzept sollte Grundlage für die Wahl der zur Beantwortung der Forschungsfragen verwendeten Modellierungstechnik sein (und nicht umgekehrt – d. h., es sollte nicht zuerst die Modellierungstechnik festgelegt werden).

4. Erst nachdem das Modellkonzept erstellt wurde, sollten die verfügbaren Daten, mit denen das Modell versehen werden könnte, ausgewählt werden. Es entspricht nicht einer guten Praxis, das Modell um die verfügbaren Daten herum zu entwerfen, da dieses Vorgehen zu restriktiv sein könnte und daraus häufig unzureichende Modelle entstehen, was zu mangelnder Sorgfalt bei der Datensammlung und anderen Problemen führen kann. Deshalb sollte die Datenverfügbarkeit, obwohl diese ggf. die Implementierung des Modellkonzepts einschränken kann, weder das Konzept noch das Einflussdiagramm bestimmen, ausgenommen die Daten stehen zu diesen in direktem Widerspruch. Stattdessen sollte das Modell das Konzept und Einflussdiagramm abbilden, von denen der Eindruck besteht, dass diese das betreffende Problem am besten erfassen, und die gesamte Funktionalität sollte in das Modell eingebaut werden. Die Analysen können dann unter eingeschränkten Datenbedingungen durchgeführt werden. Dieses ermöglicht eine flexiblere Exploration der Outcomes unter Einbezug verschiedener hypothetischer wechselseitiger Beziehungen und deren potenzieller Einflüsse.
5. Wenn nicht ausreichend Daten zur Bestückung des geplanten Modells zur Verfügung stehen, sollte eine gezielte Datensammlung für das Modell erwogen werden. Dies kann in Form von spezifischen Beobachtungsstudien erfolgen, durch den zielgerichteten Einsatz bestehender Register oder sogar durch Datenbanken, welche für andere Zwecke gesammelte Informationen beinhalten (siehe Kapitel 8.2.4).
6. Der Bewertung und möglicherweise Erhebung von Daten folgt die Entwicklung der funktionalen Zusammenhänge, die das Herzstück des Modells bilden. Die Art der zu quantifizierenden Zusammenhänge wird von der Verfügbarkeit der Daten und der gewählten Modellierungstechnik geleitet, dabei ist es entscheidend, dass diese mathematischen Beziehungen vollständig dokumentiert werden. Im Fall von *De novo*-Gleichungen (z. B. solchen, die mittels statistischer Regressionsmodelle hergeleitet werden) ist eine Dokumentation ihrer Entwicklung erforderlich.
7. Die Auswahl der Modellierungstechnik zur Strukturierung des Modells sollte zeitgleich mit der Erarbeitung der funktionalen Zusammenhänge erfolgen, da diese Aktivitäten in hohem Maße zusammenhängen. Das Hauptkriterium bei der Auswahl der Technik sollte deren Eignung zur Umsetzung des geplanten Modellkonzepts und

Einflussdiagramms sein, da nur eine angemessene Abbildung des modellierten Systems eine adäquate Beantwortung der Forschungsfragen ermöglicht. Transparenz ist ebenfalls ein wesentlicher Aspekt, aber sogar sehr komplexe Modelle können mit umfassender Dokumentation transparent sein. Übermäßige Vereinfachung von Modellen ist hingegen verhängnisvoll. Wenn das Modell das System nicht angemessen abbildet, wird es, ungeachtet seiner Transparenz, nicht brauchbar sein [9].

8. Erst nach der Validierung des Modellkonzepts, der Auswahl der Daten, der Entwicklung der verwendeten funktionalen Relationen und der Wahl der Modellierungstechnik sollte man mit der Implementierung und Programmierung des Modells fortfahren.
9. Mit der technischen Validierung sollte bereits während der Implementierung begonnen werden. Sie wird aber erst nach Abschluss der Programmierung zusammen mit der Plausibilitätskontrolle abgeschlossen.
10. Nach Abschluss der Modellvalidierung kann man mit der Analyse der relevanten therapeutischen Optionen und Zielpopulationen beginnen. Die Analysen müssen die wahrscheinlichsten Szenarien abdecken und Unsicherheiten im Bereich der Modellparameter und Modellstruktur berücksichtigen [10].
11. Der Bericht sollte alle oben beschriebenen Prozessschritte beinhalten. Bei der Beschreibung der Modellierungsmethoden besteht beträchtlicher Spielraum, sofern sie klar und vollständig ist. Die erzwungene Nutzung einer festgelegten Vorlage kann das Forscherteam dabei beeinträchtigen, das Modell auf möglichst vollständige und klare Weise zu beschreiben. Berichte müssen von einem vollständig funktionsfähigen und zugänglichen elektronischen Modell begleitet werden. Das Modell muss öffentlich zur Verfügung gestellt werden, um Vertrauen in dessen Integrität zu gewährleisten.

Die Entwicklung eines validen Modells ist ein beträchtliches Unterfangen, weshalb zur Verwendung von bereits validierten Modellen angeregt wird. Das IQWiG wird eine Datenbank mit Modellen aufbauen, die unter seiner Aufsicht entwickelt wurden und die für weitere ökonomische Evaluationen verwendet werden können. Dennoch werden

Modifikationen und Aktualisierungen der bestehenden Modelle zur Gewährleistung der Aktualität notwendig sein.

5 Einflussdiagramme

5.1 Definition

Einflussdiagramme werden zunehmend bei der Entwicklung von ökonomischen Modellen eingesetzt [11].

Ein Einflussdiagramm ist wie eine Zeichnung, die die wesentlichen Konzepte des Modells abbildet. Sie enthält die wichtigen Aspekte der zu modellierenden Erkrankung und zeigt die Beziehungen zwischen diesen auf. Trotz seines Namens ist das Einflussdiagramm nicht per se eine kausale Darstellung. Vielmehr legt es die Beziehungen zwischen den zu berücksichtigenden Aspekten offen. Hierbei handelt es sich also nicht zwangsläufig um einen Bayes'schen-Ansatz, obwohl es den Wissensstand zum Zeitpunkt des Modelldesigns wiedergibt.

5.2 Zweck

Ein Einflussdiagramm wird entwickelt, um die wesentlichen Zusammenhänge im Kern des Modells zu beschreiben und zu kommunizieren und die wichtigsten das Modell steuernden Parameter aufzuzeigen. Indem es diese explizit und visuell darlegt, ermöglicht das Diagramm allen Interessenten, auch ohne Verständnis der technischen Aspekte des Modells an sich, die Idee des Modellierers zu verstehen und Änderungsempfehlungen abzugeben. Dieses Vorgehen gewährleistet Klarheit darüber, was genau modelliert wird und auf was im Falle späterer Änderungen – z. B. im Fall einer eingeschränkten Datenlage – verzichtet werden muss, und wie viel Verzerrung in den Beziehungen dadurch entstehen kann.

Das Einflussdiagramm bildet außerdem die Basis für die Entwicklung detaillierter Flussdiagramme des Modells, die den Entwicklungsprozess unterstützen.

5.3 Komponenten

Das Einflussdiagramm muss alle wichtigen Aspekte der modellierten Erkrankung sowie deren wechselseitige Beziehungen enthalten, einschließlich:

- Patientencharakteristika
- Pathophysiologie und klinischer Verlauf der Erkrankung
- Behandlung und Management der Erkrankung

- Outcomes

Die Komponenten werden üblicherweise durch ovale Symbole dargestellt und die Beziehungen zwischen ihnen durch die Verwendung von Verbindungspfeilen, wobei der Kopf des Pfeils die Richtung der Beziehung darstellt (ohne notwendigerweise Kausalität anzuzeigen). Das Diagramm kann sehr komplex werden, das Ziel ist eine umfassende, nicht unbedingt eine einfache Darstellung. Farben und andere Kunstgriffe können helfen, die Beziehungen besser zu verdeutlichen.

In späteren Phasen der Modellierung kann das Einflussdiagramm helfen, die Datenquellen *auf einen Blick* zusammenzufassen, indem die Verbindungspfeile mit den entscheidenden gewählten Datenquellen (z. B. klinische Studie, Literatur, Datenbankanalyse) beschriftet werden, sobald diese festgelegt wurden.

Eine Erklärung der Komponenten und Verbindungen des Einflussdiagramms sollte begleitend zu dem Diagramm in einem schriftlichen Dokument bereitgestellt werden. In diesem Dokument sollte jede Komponente (d. h. ovale Form) und Verbindung erklärt sein. Darüber hinaus sollte eine Begründung für das Weglassen möglicher Verbindungen abgegeben werden.

6 Modellkonzept

6.1 Definition

Das Modellkonzept bietet die detaillierte Spezifikation der analytischen Struktur und des Ansatzes. Es ergibt sich aus dem Einflussdiagramm und stellt das beabsichtigte Design in deutlich größerer Tiefe dar.

6.2 Zweck

Die Definition des Modellkonzepts ist ein kritischer Schritt in der Modellentwicklung, sowohl im Hinblick auf die Transparenz als auch für die Bestimmung der Validität. Das Modellkonzept muss eindeutig dokumentiert und dargestellt werden. Es dient dem Zweck sicherzustellen, dass das modellierte Indikationsgebiet und die Effekte der Interventionen genau verstanden wurden und dass alle Annahmen bezüglich der Beziehungen zwischen den Modellkomponenten explizit genannt sind. Da auch die differenziertesten Modelle Vereinfachungen der Realität mit erforderlichen Annahmen und Einschränkungen in Bezug auf eingeschlossene Inhalte [10,12] darstellen, kann die Validität und Nützlichkeit des Modells nur dann richtig verstanden werden, wenn das Modellkonzept klar spezifiziert wird.

6.3 Komponenten

Die Spezifizierung des Modellkonzepts sollte folgende Aspekte mit berücksichtigen:

- Welche klinischen Ereignisse sollten mit einbezogen werden und wodurch wird deren Ausprägung beeinflusst?
 - Können diese Ereignisse wiederholt auftreten und wodurch wird das Wiederholungsrisiko beeinflusst?
 - Beeinflusst das Erleben eines Ereignisses die Wahrscheinlichkeit oder den Zeitpunkt von nachfolgenden Ereignissen?
 - Muss der Schweregrad eines Ereignisses berücksichtigt werden?
- Welche Gesundheitszustände müssen berücksichtigt werden?

- Beschreiben die Gesundheitszustände hinreichend das klinische Spektrum der Erkrankung?
- Sind sie ausreichend, um Patienten hinsichtlich Gesundheitsstatus, gesundheitsbezogener Lebensqualität und Kosten zu unterscheiden?
- Erfassen sie hinreichend die Geschichte und Erfahrung der Patienten?
- Wie beeinflusst der Nutzen der Interventionen, wie er vom IQWiG festgelegt wird, den Verlauf der Erkrankung?
 - Werden Ereignisse hinausgezögert oder gänzlich vermieden und / oder in ihrem Schweregrad verringert?
 - Wie sehen der zugrunde liegende Mechanismus und die Annahmen aus, die den Einfluss der Intervention auf den natürlichen Verlauf der Erkrankung bestimmen?
 - Welche Annahmen über die Langzeitwirkung der Interventionen und der Folgen einer Behandlungsverzögerung oder -aussetzung werden gemacht?
- Wie beeinflusst das Verhalten der medizinischen Versorger oder Patienten die Outcomes?
 - Ist die Compliance der professionellen Gesundheitsversorger oder der Patienten ein wichtiges Kriterium?
 - Wie werden zukünftige medizinische Entscheidungen zum Patientenmanagement vom Ansprechen auf die Behandlung und / oder von der Compliance beeinflusst?
 - Ist die Interaktion zwischen Patienten relevant, wie beispielsweise im Fall von akuten Infektionserkrankungen?

- Wie werden die Kosten durch das Patientenmanagement, die Patientenhistorie, das Patientenverhalten und den Zeitpunkt bzw. den Schweregrad von Komplikationen beeinflusst?

6.4 Validierung

Modelle sollten „*consistent both with a coherent theory of the health condition being modeled and with available evidence regarding causal linkages between variables*“ [5] sein. Deshalb kann die Fertigstellung des zentralen Modellkonzepts nicht ohne die Validierung fortgesetzt werden.

Zur Gewährleistung der Validierung des Modellkonzepts ist es wichtig, eine vollständige Dokumentation der zur Definition des Modellkonzepts verwendeten Information und der hypothetischen Verbindungen mit einzuschließen. Auch ist es hilfreich, das Konzept mit demjenigen relevanter existierender Modelle zu vergleichen (Kreuzvalidierung des Modells). Dieser Vergleich sollte eine Beschreibung und Begründung für jede Abweichung von Konzepten, die üblicherweise bei der betreffenden Erkrankung und in dem betreffenden Indikationsbereich zur Anwendung kommen, beinhalten.

Übereinstimmung mit publizierten ökonomischen Evaluationen ist jedoch angesichts der unterschiedlichen Qualität von publizierten Studien und des Fehles von einer Übereinstimmung, was ein „gutes Modell“ [13] ausmacht, unzureichend. Modelle müssen für Personen, die sich mit der Erkrankung auskennen, Sinn ergeben [14]. Zur Validierung des Konzepts müssen deshalb anerkannte Experten einbezogen werden, die sich mit den modellierten Indikationsgebieten auskennen, ohne sich dabei zwangsläufig auf Kliniker zu beschränken. Relevante Expertise betrifft die Bereiche Epidemiologie, Statistik, Lebensqualität, Public Health und Ökonomie.

7 Erforderliche Modelleigenschaften

Die Kostenschätzung unter Verwendung von Modellen muss ein Minimum an Anforderungen erfüllen, um für die Zwecke des IQWiG geeignet zu sein. Dieses Kapitel behandelt diese Anforderungen, wobei anzumerken ist, dass die Modellqualität nicht nur auf Basis der Modelleigenschaften vorgegeben wird, sondern Peer-Review-Verfahren und vollständig zugängliche Modelle einen entscheidenden Bestandteil all dieser vom IQWiG erwogenen Kostenschätzungen darstellen.

7.1 Hinreichende Detailliertheit

Wie in Kapitel 9 detailliert dargestellt, müssen die IQWiG-Modelle zur Beantwortung der Fragen hinsichtlich der mit einer betreffenden Gesundheitstechnologie zusammenhängenden Kosten über ein ausreichendes Ausmaß an Detailliertheit verfügen. Dies erfordert folgende Modelleigenschaften:

- Erfassung von allen Krankheits- und Behandlungsaspekten, die mit einer gewissen Wahrscheinlichkeit einen bedeutsamen Einfluss auf die Kosten haben,
- Einbezug von Nutzen und Schaden durch Interventionen,
- Berücksichtigung der relevanten Heterogenität in der Zielpopulation, des natürlichen Verlaufs der Erkrankung und der Interventionseffekte,
- Abbildung von praxisüblichen Behandlungsmustern in Deutschland,
- Einbezug der Effekte einer nicht optimalen Behandlungscompliance,
- Möglichkeit, die Outcomes über den Zeitverlauf hinweg darzustellen sowie pro Kostenkategorie, anstatt nur als über ein festes Zeitintervall aggregierte Kosten und
- Möglichkeit zum Einbezug verschiedener demografischer, epidemiologischer, Kosten- und Behandlungsmusterdaten, um für verschiedene Regionen in Deutschland relevante Ergebnisse erzeugen zu können.

Wo keine ausreichenden Daten zur Verfügung stehen, um diese Merkmale zuverlässig abzudecken, ist es wichtig, den möglichen Einfluss auf die Kostenschätzungen durch angemessene Wertebereiche der fehlenden Daten zu explorieren.

7.2 Perspektive

Als Perspektive der Untersuchung sollte die Perspektive der Gemeinschaft der in der gesetzlichen Krankenversicherung versicherten Bürger (GKV-Versichertengemeinschaft) gewählt werden. Diese ist nicht identisch mit der Perspektive der gesetzlichen Krankenversicherung selbst, da sie Ressourcen der Patienten und Gesellschaft mit einbezieht.

Sofern möglich, sollten Ergebnisse zu Kosten sowohl in aggregierter als auch in vollständig disaggregierter Form zumindest im Hinblick auf die folgenden Kategorien (welche sich gegenseitig nicht völlig ausschließen) berichtet werden:

- Evaluierte Gesundheitstechnologien (z. B. Medikamentenkosten)
- Andere Aktivitäten, die sich direkt der evaluierten Gesundheitstechnologie zuordnen lassen (Monitoring der Behandlung von Nebenwirkungen)
- Krankenhauseinweisungen
- Notaufnahme
- Besuche bei niedergelassenen Ärzten
- Andere Fachleute
- Ausstattung oder Apparate
- Sonstiges (siehe Technische Anhänge – Kostenbestimmung, Kapitel 2.2)

7.3 Zeithorizont

Der Zeithorizont der Modellierung sollte die Erfassung aller Kosten erlauben, die durch die betreffende Gesundheitstechnologie beeinflusst werden. Dieser kann recht lang sein, z. B. lebenslang im Falle chronischer Erkrankungen. Abhängig von der evaluierten Intervention kann der Zeithorizont sogar über die Lebenszeit der direkt betroffenen Personen hinausgehen,

z. B. bei der Evaluation von Impfungen. Deshalb muss eine nachvollziehbare Dauer gewählt werden, die die für die Erstattungsentscheidung wichtigsten Kosten abdeckt. Sollte ein kürzerer Zeithorizont herangezogen werden, muss diese Entscheidung explizit begründet werden.

7.4 Behandlungsdauer

Die Behandlungsdauer sollte unter Verwendung von mindestens zweier Szenarien evaluiert werden:

- Behandlungen dauern so lange wie in den Studien, in denen der Nutzen der Behandlung untersucht wurde, und
- die Behandlungsdauer umfasst den für die Erstattungsentscheidung relevanten Zeitraum (z.B. die übliche Praxis in Deutschland).

Die Behandlungsdauer muss zusätzlich eine vorzeitige Behandlungsunterbrechung mit berücksichtigen, sei es aufgrund des Patientenverhaltens, der Nebenwirkungen, des ärztlichen Rats oder aufgrund von fehlenden Therapieeffekten.

7.5 Diskontierung

Alle in dem Modell evaluierten Kosten müssen in angemessen diskontierter Form berichtet werden. Der Leser sei zur Diskontierung von Kosten auf Kapitel 3.2.5.2 sowie auf den Technischen Anhang zur Kostenbestimmung verwiesen. Sensitivitätsanalysen sollten eine weite Spanne an Diskontierungsraten abdecken, um die Auswirkung des Diskontierungseffekts auf die Nettokosten der Intervention zu untersuchen. Diskontierungsraten werden vom IQWiG festgelegt und sofern erforderlich, aktualisiert.

8 Datenbewertung

Die verfügbaren Daten sollten systematisch erforscht und ihre Qualität und Relevanz zur Unterstützung der Modellinputs und Risikofunktionen muss untersucht werden. Die Qualitätskriterien, die von Epidemiologen für jeden Evidenzgrad entwickelt wurden, werden im Allgemeinen von Experten aus diesem Bereich als ebenso anwendbar für die Bewertung von Daten, die ökonomische Studien untermauern, angesehen [15], aber dies ist nicht korrekt. Daten aus gut durchgeführten randomisierten kontrollierten Studien werden also als empfehlenswerter angesehen als solche aus nicht-randomisierten Studien. Für Kosteninformationen sind sie jedoch eher von geringer Relevanz, wenn sie nicht mit dem üblichen Vorgehen in der Praxis übereinstimmen [15]. Ebenso werden kontrollierte prospektive Studien gegenüber retrospektiven Studien als überlegen angesehen, aber für die Kosteninformationen können auch eine retrospektiv gebildete Datenbankkohorte oder ein Fall-Kontroll-Ansatz nützlich sein.

Üblicherweise werden Daten aus verschiedenen Datenbanken gewonnen, einschließlich aus Studien zu relativen Behandlungseffekten, Kohortenstudien zur Untersuchung von Parametern und Risikofaktoren für den natürlichen Krankheitsverlauf und die damit verbundene Lebenserwartung, sektorenübergreifenden Fragebögen zur Erfassung von Lebensqualitätsdaten, Ressourcenverbrauch und Kosten, Registern, Prozessdatenbanken oder kompilierten Statistiken. Auf Expertenmeinungen basierende Annahmen sollten vermieden werden, da diese selten exakt genug für eine Verwendung bei fehlender Verfügbarkeit anderer Datenquellen sind. Ein wichtiges Kriterium für Studien ist die Übertragbarkeit der Ergebnisse auf den deutschen Kontext.

Die Begründung für die Auswahl bestimmter Datenquellen sollte in den Dokumenten enthalten sein (beispielsweise die verwendeten Selektionskriterien, die externe Validität), insbesondere für die Parameter, die die Modellergebnisse wesentlich beeinflussen. Details der Datenverarbeitungsmethoden zur Generierung der benötigten Inputs, Funktionen und Verteilungen der Parameter müssen ebenfalls mit dargestellt werden [13,16,17].

8.1 Ziele

Zur Bewertung der verfügbaren Daten muss deren Qualität und Relevanz für die spezifischen Modellierungsziele überprüft werden. Außerdem sind die Datenlücken zu identifizieren, bei denen sinnvolle Annahmen getroffen werden können, und eine geeignete Methode zur Bearbeitung der Daten ist anzuwenden, um reliable Schätzer zu erhalten.

8.2 Potenzielle Datenquellen

8.2.1 Klinische Studien

Die vom IQWiG bewerteten klinischen Endpunkte sind Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und valide Surrogate. Um nach IQWiG-Kriterien verwendbar zu sein, muss bei Surrogaten überzeugend anhand von Studien dargelegt werden, dass Veränderungen in den klinischen Messwerten mit Änderungen in patientenrelevanten Endpunkten korrespondieren. Kriterien, denen ein valides Surrogat genügen muss, sind im allgemeinen Methodenpapier des IQWiG beschrieben. Es ist Gegenstand der vorgeschalteten Nutzenbewertung, valide Surrogate zu identifizieren und, sofern solche vorliegen, den Effekt der zu betrachtenden Interventionen auf diese Surrogate zu evaluieren. Derartige Surrogate sind dann eine Option für ökonomische Evaluationen mittels Modellierung. Die patientenrelevanten Nutzeneffekte diagnostischer und therapeutischer Interventionen werden in der vorgeschalteten Nutzenbewertung des IQWiG zur Verfügung gestellt und in kontrollierten klinischen Studien ermittelt. Valide Modelle können ggf. Vorhersagen darüber treffen, wie sich die in klinischen Studien ermittelten Nutzeneffekte unter anderen Bedingungen oder Zeitperspektiven darstellen würden. Modelle können jedoch in keinem Fall neue Nutzeneffekte generieren.

Abhängig von der Qualität, Relevanz und Verfügbarkeit der Daten können im Modell Inputwerte verwendet werden, die auf der relativen Wirksamkeit aus einer Studie basieren, oder aus gepoolten Datensätzen bzw. einer Meta-Analyse berichteter Studienergebnisse stammen. Im Allgemeinen sind zum Zeitpunkt der Evaluation eines neuen Produktes im Hinblick auf die Erstattung nur relativ kleine klinische Wirksamkeitsstudien (klinische Prüfungen) mit kurzem Zeithorizont fertig gestellt. Die Dauer dieser klinischen Wirksamkeitsstudien ist häufig unzureichend zur Aufdeckung von signifikanten Unterschieden bezüglich Kosteninformationen. Dieser Sachverhalt bildet eine der

wesentlichen Limitationen bei ökonomischen Evaluationen, da diese das Versorgungsgeschehen abbilden müssen – was sich über größere Zeithorizonte erstrecken kann [18-20].

Daten aus Studien sollten nach Überprüfung ihrer Relevanz extrahiert werden, insbesondere im Hinblick auf die Wahl der Vergleichsstrategien, gemessenen Outcomes, Surrogate vs. harte Endpunkte, Länge des Follow-ups, Beteiligung von Ländern in der Studie, Ausschlusskriterien, die zu einer Unterrepräsentierung von Patientensubgruppen führen, den durch das Protokoll verursachten Ressourcenverbrauch und die Compliance der Patienten oder Ärzte [18-20]. Die erfolgten Schritte für die Umwandlung der extrahierten Daten in eine für die Einbindung in das Modell geeignete Form sollten beschrieben und jede Anpassung oder Kalibrierung transparent gemacht werden.

Die in klinischen Studien gemessenen Outcomes werden im Modell mit dem Ressourcenverbrauch verknüpft, wie z. B. Krankenhauseinweisungen, Behandlungsverfahren oder andere Dienstleistungen. In vielen Fällen wird erforderlich sein, die Studiendaten mit einschlägigen Informationen über das deutsche Gesundheitssystem unter Verwendung von Datensammlungen aus nichtrandomisierten Studien zu ergänzen. Die übliche Versorgungspraxis spiegelt die örtlichen Erstattungsregeln und den Zugang zu gewissen Leistungen in jedem Land wider. So wird beispielsweise der Zugang zu Pflegeeinrichtungen oder die Zahlungen an pflegende Familienangehörige den Ort der Pflege beeinflussen. Viele Studien werden international durchgeführt, und Strukturen des Ressourcenverbrauchs, Ort der Pflege sowie epidemiologische Parameter, die die Kosten beeinflussen, variieren beträchtlich zwischen den Ländern und können deshalb nicht ohne Begründung direkt auf Deutschland übertragen werden [21,22]. Studienprotokolle können den Verbrauch von Ressourcen erforderlich machen, die in der normalen Versorgung nicht anfallen würden. So können andere Datenquellen zur Identifikation des Ressourcenverbrauchs in der klinischen Praxis hilfreich sein, da es nicht angebracht ist, die durch das Protokoll erzeugten Kosten in die Kostenschätzung mit einzubeziehen. Alle Details des Ressourcenverbrauchs, die Studien entstammen, sollten im Hinblick auf ihre Relevanz für das deutsche Gesundheitssystem überprüft und im eingereichten Bericht gesondert beschrieben werden, da dieses großen Einfluss auf die letztendlichen Schlussfolgerungen der Studie haben kann [23].

Meta-Analysen von Studien werden häufig zur Begründung der relativen Behandlungseffekte eingesetzt, und Details der in der Meta-Analyse verwendeten Methoden sollten berichtet werden. Head-to-head-Vergleiche von Interventionen sind nicht immer verfügbar, wodurch indirekte Vergleiche erforderlich werden können, um die zu bewertenden Behandlungen mit einzubeziehen. Es sollten geeignete Methoden zur Ableitung dieser indirekten Schätzer der Behandlungswirksamkeit verwendet werden [24-26].

Die transparente und vollständige Dokumentation der zur Gewinnung der benötigten Inputs vorgenommenen Datenanalyse oder Datenbearbeitung wird als qualitätssichernde Maßnahme und Gütekriterium der Modellierung betrachtet.

8.2.2 Epidemiologische Studien

Daten aus Beobachtungsstudien sind häufig notwendig, um Daten aus klinischen Studien zu erweitern oder zu extrapolieren und eine Modellierung über den Zeitraum der klinischen Studie hinaus zu ermöglichen. Bei vielen Erkrankungen vergehen Jahre oder Jahrzehnte bis zum Erreichen der finalen Gesundheitsoutcomes. In solchen Fällen ist es üblich, intermediäre klinische Endpunkte zu untersuchen, welche als Surrogat für die finalen und für Patienten tatsächlich relevanten Endpunkte, wie Langzeitmorbidity und Mortalität, dienen [7]. Bei diesen Surrogaten muss anhand durchgeführter Studien überzeugend nachgewiesen werden, dass Veränderungen in den intermediären klinischen Endpunkten mit Änderungen in patientenrelevanten Endpunkten korrespondieren.

Die Interpretation jeder prädiktiven Funktion und die Validität der Übertragung feststehender Risikofunktionen auf die modellierte Population sowie jegliche vorgenommene Anpassungen müssen erklärt und begründet werden. Gründe für die Verwendung von Beobachtungsdaten sollten dargelegt werden. Details der verwendeten statistischen Methoden und aller durchgeführten Analysen zur Entwicklung angemessener Schätzer für die Integration in das Modell müssen dargestellt werden. Die Validität der Annahmen, dass diese Funktionen oder Schätzer für die in einem Modell für den deutschen Kontext betrachtete Population verwendet werden können, muss beurteilt und jede Modifikation erläutert werden.

8.2.3 Prozessdatenbanken

Medizinische Abrechnungsdaten können Details von Ressourcenverbräuchen in der Gesundheitsversorgung für eine große Anzahl von Patienten in der tatsächlichen Praxis bieten. Daher sind sie nützliche Datenquellen, wenn die für das Modell relevante Zielpopulation durch die in der Datenbank verwendeten diagnostischen Codes ausreichend zu definieren ist. Problematisch ist, wenn die Populationsdefinition klinische Befunde, Laborergebnisse, andere Patientencharakteristika als Alter und Geschlecht oder sonstige Items einschließt, die üblicherweise nicht in solchen Datenbanken codiert werden. Auch Abrechnungsdaten von Arzneimitteln können Informationen zu in der Realität üblichen Behandlungspfaden, Wechselverhalten, Compliance (zumindest bezüglich Arzneimittelbeschaffung, wenn auch nicht über die tatsächliche Arzneimitteleinnahme) und Persistenz liefern.

Die Relevanz einer vorhandenen Datenbank für ein bestimmtes Modell muss bewertet werden, da viele Faktoren die Verschreibungspraxis beeinflussen können, einschließlich lokaler Erstattungsbedingungen, Markteinführungszeitpunkten und sogar des Prozesses der Veröffentlichung von Daten, da hier eine Verzögerung entstehen kann, bis Informationen für die Analyse verfügbar sind. Obwohl eine vorhandene Datenbank nützlich sein kann, den Ressourcenverbrauch bei bestimmten Krankheitszuständen zu quantifizieren, ist bei vergleichenden Analysen große Vorsicht geboten, da es zu erheblichen Verzerrungen kommen kann, wenn die Patienten den verglichenen Behandlungen nicht randomisiert zugeordnet wurden [16].

8.2.4 Register

Patientenregisterstudien sind Beobachtungsstudien, bei denen ein gekürzter Datensatz von einer großen Anzahl an Patienten, die unter Routinebedingungen behandelt werden, systematisch erfasst wird. Da die Daten gewöhnlich im Rahmen der Regelversorgung erhoben werden, tendieren sie eher dazu, die aktuelle klinische Praxis genau widerzuspiegeln, und bieten eine gute Informationsquelle. Im Gegensatz zu klinischen Studien neigen Registerstudien nicht in gleichem Maße dazu, die Zugangskriterien oder Behandlungsoptionen strikt vorzugeben.

Ein gut konzipiertes Register kann eine wertvolle Datenquelle für ein Modell sein. Es kann zur Entwicklung vieler der benötigten prädiktiven Funktionen, die die Assoziation zwischen klinischen Ereignissen und Ressourcenverbrauch, Versorgungsqualität, Arbeitsfehltagen etc. beschreiben, beitragen.

Die Bewertungen des Nutzens von Registerdaten sollte eine Überprüfung der Relevanz unter Berücksichtigung der Zielsetzung des Registers, seines Designs, seiner Patientenpopulation(en) und seiner Durchführung umfassen, um die Qualität und Vollständigkeit der erhobenen Daten zu bestätigen (z. B. durch Überprüfen des Anteils an fehlenden Werten wichtiger Variablen und durch Erfassen, wie diese behandelt und berichtet wurden) [27]. Die Analysen, die durchgeführt wurden, um die fürs Modell geeigneten Formen von Funktionen und Schätzern zu entwickeln, müssen detailliert ausgeführt werden. Darüber hinaus muss im Sinne der Qualitätssicherung ein angemessener Zugang zum Register gewährleistet werden.

8.2.5 Kompilierte Statistiken

Kompilierte Statistiken sind Gesundheitsstatistiken, die von Regierungen aus Zensus- oder Survey- Daten kompiliert wurden. Diese Quellen können relevante Daten zur Demografie in der Population, wie Alter, Geschlecht und gesundheitsrelevantes Verhalten sowie Gewicht etc. liefern [28]. Kompilierte Statistiken werden häufig genutzt, um Mortalität aus anderen Gründen als den für das Modell relevanten, abzuschätzen (z. B. die Human Mortality Datenbank [29]). Jede im Modell verwendete Quelle kompilierter Daten muss genannt und in Hinblick auf Qualität und Relevanz bewertet werden, um sicherzustellen, dass diese angemessen, ausreichend aktuell und vollständig ist.

8.2.6 Expertenmeinung

Expertenmeinungen, Fokusgruppen und Konsensus-Panels können in sehr begrenzter Weise in Modellierungsstudien verwendet werden: zur Beeinflussung der Modellstruktur und zur Unterstützung der in Einflussdiagrammen und Modellkonzept getroffenen Annahmen. Es wurde berichtet, dass Expertenmeinungen auch in Situationen genutzt werden können, in denen Grenzen in Bezug auf verfügbare Daten sowie Ressourcen für Forschung zur Durchführung zusätzlicher Studien bestehen, um präzisere Schätzer zu erhalten [13,30]. Unter diesen Umständen sei ein pragmatischer Ansatz notwendig und die verfügbare Evidenz könne

zu diesem Zeitpunkt durch sachkundige Meinungen ersetzt werden [13,17,30]. Andere führen an, dass die Durchführung einer ökonomischen Analyse mit der Nutzung von Expertenmeinungen von Gutachtern als methodischer Fehler angesehen wird [31].

Für tatsächliche Inputwerte und mehr noch für funktionale Beziehungen ist diese Informationsquelle nahezu niemals angemessen [16,32-34]. Es wird empfohlen, dass bei Modellierungsstudien, die vom IQWiG durchgeführt werden, die Verwendung von Expertenmeinungen minimiert wird und auf der Generierung aktueller Daten (wenn möglich, während der Bewertung) zur Unterstützung von Schätzungen bestanden wird.

Expertenmeinungen können als wertvoll erachtet werden im Hinblick darauf, was beim Management von Patienten getan wird (nicht wie häufig) und können den Transfer erhobener Daten zu Ressourcenverbräuchen zwischen Ländern erleichtern, wenn entsprechende Leistungen nicht verfügbar sind. Die Rechtfertigung für die Nutzung von Expertenmeinungen und eine Beschreibung der verwendeten Methoden zur Eruierung dieser Meinungen müssen im Modellbericht aufgeführt [32] und die Variabilität in den erhobenen Meinungen wiedergegeben werden [13,32,34,35].

9 Funktionale Beziehungen

9.1 Definition

Eine funktionale Beziehung ist eine Gleichung, welche eine Response-Variable ihren Determinanten, das heißt Variablen, die mit der Response-Variablen korreliert oder assoziiert sind, zuordnet. Unter den Kovariablen finden sich üblicherweise Patientencharakteristika und Zeit. Diese für das Modell benötigten funktionalen Beziehungen werden unter Einsatz geeigneter statistischer Techniken aus den verwendeten Datenquellen abgeleitet. Alternativ können publizierte Gleichungen genutzt werden, sofern sie für die modellierte Population passen. Dieses Kapitel beschreibt die Arten von Beziehungen, die betrachtet werden sollten, und geeignete statistische Techniken zu ihrer Herleitung.

9.2 Zweck

Die angegebenen Funktionen sollten die im Einflussdiagramm des Modells postulierten Beziehungen widerspiegeln. D. h. mithilfe dieser Gleichungen soll sichergestellt werden, dass assoziierte Parameter angemessen verknüpft sind, sodass eine Änderung im einen Parameter eine Änderung im anderen Parameter auslöst (hier als „Effekt“ bezeichnet, ohne notwendigerweise Kausalität zu implizieren). Die Vernachlässigung von Beziehungen zwischen Parametern oder, äquivalent, das Auslassen wichtiger Faktoren in den im Modell angegebenen funktionalen Beziehungen führt zu einer Mittelung des Effekts (also der Response-Variablen) über mehrere Stufen oder Werte der ausgelassenen Eingangsvariablen. Der Einschluss von Variablen, die zu Effekt- und Kostenparametern korreliert sind, führt zu genaueren Schätzungen und reduziert dadurch die Unsicherheit der Ergebnisse ökonomischer Schätzungen [36,37].

9.3 Allgemeine Konzepte

Funktionale Beziehungen sind besonders wichtig bei simulationsbasierten ökonomischen Evaluationen, da diese die Simulation kontrollieren. Multiple Regressionstechniken und andere Methoden können angewandt werden, um diese funktionalen Beziehungen zu erzeugen [36,38,39].

Funktionale Beziehungen sollten bestimmt werden, um Effekte und Kostenparameter auf Patientencharakteristika, Behandlung, Zeit und andere relevante Faktoren zu beziehen. Im einfachsten Fall würden die Effektparameter nur auf die einzige interessierende Determinante bezogen: Für eine bestimmte Intervention würden diese einen „mittleren“ Effekt für die Intervention und jede in Betracht gezogene Alternative liefern. Werden auf der anderen Seite Patientencharakteristika, die mit dem Effekt assoziiert sind, einbezogen, erlaubt dies patientenspezifische Vorhersagen bei Änderung in den betroffenen Parametern. Dieses System erlaubt auch die Integration von Subgruppeneffekten durch den Einbezug von Interaktionstermen. Der Ressourcenverbrauch kann in ähnlicher Weise modelliert werden, um die Abhängigkeit vom Patiententyp zu erfassen; so können zum Beispiel ältere Patienten und solche mit einer schweren Erkrankung höhere Kosten verursachen. Die statistischen Eigenschaften der Daten zum Ressourcenverbrauch weisen erhebliche Schwierigkeiten auf, deren Behandlung besondere Sorgfalt erfordert [40].

Wenn verschiedene Effekte tendenziell korrelieren, sollten sie durch funktionale Beziehungen verknüpft werden. Zum Beispiel kann der Effekt eines Patientencharakteristikums wie die Erreichbarkeit eines Pflegers in Bezug auf die Krankenhausaufenthalte und die Nutzung eines ambulanten Pflegedienstes betrachtet werden; ein großer Rückgang bei den Krankenhausaufenthalten geht wahrscheinlich mit einer starken Zunahme der ambulanten Pflegedienste und vice versa einher. Um diese zu verbinden, können zum Beispiel Responses bezogen auf ein Outcome als Kovariable in die funktionale Beziehung für den anderen Endpunkt einbezogen werden.

Auch in ökonomischen Modellen, in denen die Datenquelle statistische Verteilungen zur Verfügung stellt, aus denen Patientenmerkmale gezogen werden können, müssen Korrelationen zwischen den Merkmalen betrachtet werden. Zum Beispiel ist das Gewicht mit dem Geschlecht des Patienten assoziiert; Risikofaktoren neigen dazu zu korrelieren, die Schwere der Erkrankung kann aus den Patientencharakteristika vorhersagbar sein etc. Die Erfassung von Korrelationen zwischen allen Charakteristika kann schwierig sein; stattdessen kann ein Bootstrapping der Patientenprofile aus der Datenquelle verlässlicher sein, da dies bei einer entsprechend großen und variierenden Stichprobe sicherstellt, dass alle Beziehungen simultan abgedeckt werden.

In Situationen, in denen Effekte longitudinal oder in Form von *Time-to-event*- Daten gemessen werden, müssen funktionale Beziehungen wiedergeben, wie das Outcome sich *per se* über die Zeit ändert; zusätzlich ist es wichtig zu berücksichtigen, ob die Effekte im Laufe der Zeit variieren. Im Zusammenhang mit Überlebenszeitanalysen (*Time-to-event*-Analysen) ist eins der am häufigsten verwendeten Responsemaße das Risiko eines Endpunktes (Tod oder einige andere Ereignisse); die Änderungsmuster des Risikos im Laufe der Zeit werden durch die Art der statistischen Verteilung, die den Ereigniszeiten zugeordnet wird, bestimmt. Bei longitudinalen Analysen, die andere Aspekte als den Eintritt von Ereignissen (z. B. Änderungen des Blutzuckers) berücksichtigen, sollten Änderungen der Response im Laufe der Zeit durch die Verwendung einer Vielfalt an Funktionsarten der Zeitparameter untersucht werden. In beiden Fällen sind die Effekte eines anderen Faktors an verschiedenen Zeitpunkten zu evaluieren; Änderungen der Effekte sollten bewertet und in die funktionalen Beziehungen zwischen Endpunkt und Kovariaten einbezogen werden. Wenn Daten an mehreren Standorten erhoben werden, ist es wichtig, den Einfluss des Standorts auf die Effekte und Kostenparameter zu untersuchen [41,42]. Potenzielle Verzerrung muss durch die Wahl einer geeigneten Technik kontrolliert werden [42-45].

Kontinuierliche Zielgrößenvariablen wie Blutdruck, Gewicht etc. können mit linearen Regressionstechniken modelliert werden, solange deren Verteilung näherungsweise einer Normalverteilung entspricht. Dies gilt zum Beispiel nicht für Kostendaten, welche zu einer sehr schiefen Verteilung neigen. Logarithmische Transformationen oder alternative Modelle basierend auf Gamma-Verteilungen sollten in Erwägung gezogen werden [38,42,46]. Wenn die Schiefe in den Kosten durch eine übermäßige Zahl von Patienten ohne Kosten verursacht wird, ist zweistufiges Vorgehen angebracht, bei dem zuerst die Wahrscheinlichkeit der Verursachung von Kosten modelliert wird, gefolgt von einer Analyse der Höhe der Kosten unter den Patienten, für die überhaupt Kosten anfallen [47].

In einigen Fällen benötigen ökonomische Modelle eine Vorhersage des Zeitintervalls eines Ereignisses, um zum Beispiel die Dauer des Einsatzes einer Behandlung abzuschätzen. In diesen Situationen sind nichtparametrische oder parametrische Analysen, wie beispielsweise die Überlebenszeitanalyse, notwendig, um Änderungen des Risikos eines Ereignisses im Zeitverlauf zu beschreiben. Dies erfolgt durch die Bestimmung einer statistischen Verteilung für die Ereigniszeiten; die am häufigsten verwendeten Verteilungen sind

Exponentialverteilungen sowie die Weibull- und die Gompertz-Verteilung [48]. Die Exponentialverteilung setzt voraus, dass das Ereignis mit einer konstanten Rate oder einem konstanten Risiko auftritt. Da die meisten Risiken in der Medizin im zeitlichen Verlauf nicht konstant sind, ist diese Verteilung selten von Nutzen. Die Weibull-Verteilung unterstellt, dass die Risiken über die Zeit hinweg einen monoton steigenden oder fallenden Verlauf haben (exponentiell als Spezialfall). Die Risiken in der Gompertz-Verteilung steigen oder fallen ebenfalls monoton, aber mit einer schnelleren Rate als in der Weibull-Verteilung. Die log-normalen oder log-logistischen Verteilungen sollten ebenfalls beachtet werden [48]. All diese Verteilungen, außer der exponentiellen, weisen einen Form- und einen Skalierungs-Parameter auf. Beide Parameter können entsprechend Patientencharakteristika und anderer Determinanten variiert werden.

Evidenz, welche die Validität der in dem ökonomischen Modell verwendeten funktionalen Beziehungen unterstützt, sollte dargestellt werden. Dies sollte die Beschreibung aller Annahmen und getroffenen Entscheidungen bei der Entwicklung der Gleichungen einschließen. Darüber hinaus sollten ausreichende Informationen zu den Datenquellen gegeben werden, um zu zeigen, dass die Daten adäquat für die durchgeführten Analysen waren.

Verschiedene Strategien können angewandt werden, um bei der Entwicklung einer prädiktiven Gleichung Variablen zu selektieren. Diese beinhalten automatische Selektionsprozeduren (Vorwärts-, Rückwärts-, schrittweise Selektion) oder manuelle Selektion von Variablen, basierend auf statistischer Signifikanz, oder frühere klinische Erkenntnisse über die zu untersuchenden Beziehungen. Das IQWiG empfiehlt in Übereinstimmung mit seinen Allgemeinen Methoden die Rückwärts-Selektion [49]. Da ein vollständiges Fehlen klinischer Erkenntnisse über die Beziehungen extrem selten ist, wird empfohlen, das vorhandene klinische Wissen bei jedem Modellbildungsprozess zu berücksichtigen. In jedem Fall sollte der verwendete Ansatz zur Auswahl der Variablen für jede Gleichung beschrieben und begründet werden.

Es sollten auch explizite Informationen über die verwendeten Kriterien für die Wahl der Art der Gleichung geboten werden. Zum Beispiel kann, im Falle einer linearen Gleichung die Annahme zugrunde liegen, dass die Kriterien die erklärte Varianz maximieren sollten. Bei logistischen und Überlebenszeit-Analysen können prädiktive Funktionen hinsichtlich ihrer

Diskriminationskraft evaluiert werden; Maße wie der *c-Index* oder *Receiver Operating Characteristic*-(ROC)-Kurven können verwendet werden [50].

Die bei der Erstellung der funktionalen Beziehungen benutzten Annahmen sollten zusammen mit den verwendeten Ansätzen zur Sicherstellung ihrer Gültigkeit dargestellt werden. Die Güte der Modellanpassung („Goodness-of-Fit“) der gewählten Regressionsmodelle sollte untersucht und anhand üblicher Verfahren beschrieben werden [50,51]. Wenn kausale Funktionen entwickelt werden, sollten Annahmen zur Existenz und Richtung kausaler Beziehungen anhand vom Kausalgraphen dargestellt werden [52].

10 Modellierungstechniken

Verschiedene Modellierungstechniken sind verfügbar und wurden bereits in gesundheitsökonomischen Evaluationen eingesetzt wie z. B. das Entscheidungsbaumverfahren, Markov-Modelle, die Diskrete Ereignissimulation (DES), Transmissionsmodelle, Agentenbasierte Modelle und viele andere [3,6,12,53-57]. Es ist besonders wichtig, dass eine Festlegung auf eine bestimmte Modellierungstechnik nicht im Voraus geschieht. Daher hat das IQWiG keine A-priori-Präferenz für eine besondere Modellierungstechnik. Die Wahl der geeigneten Modellierungstechnik hängt vielmehr von der wissenschaftlichen Fragestellung ab, die dem IQWiG vom Gemeinsamen Bundesausschuss (G-BA) im Rahmen des Auftrags gestellt wird, sowie von den Charakteristika der zu bewertenden Technologie, Erkrankung und den Rahmenbedingungen. Der geeignete Prozess zur Entwicklung effizienter Modelle, die alle relevanten Bewertungsprobleme berücksichtigen, erfordert vorab, das Modell vollständig zu verstehen und zu konstruieren und die verfügbaren Daten und funktionalen Beziehungen zu erheben. Erst dann sollte eine Modellierungstechnik gewählt werden.

Letztendlich wird die Auswahl der Modellierungstechnik durch eine Vielzahl von Faktoren bestimmt, jedoch sollte das Leitprinzip sein, dass das ökonomische Modell so differenziert und komplex sein sollte, wie es für die adäquate Beantwortung der gestellten Forschungsfragen erforderlich ist. Vereinfachungen des Designs zum Zweck der Einsparung von Entwicklungs- oder Analysezeit oder um Nicht-experten das Verständnis des Modells zu erleichtern, sind nicht akzeptabel, falls die Modelle für reale Entscheidungsfindung benutzt werden sollen. Genauso sollte die Wahl der Modellierungstechnik nicht aus Gewohnheit erfolgen. Mit einer adäquaten Dokumentation können selbst komplexe Modelle transparent dargestellt werden und Computer mit zunehmender Rechenleistung, sowie Techniken der Varianzreduktion führen dazu, dass die Zeit und technischen Anforderungen zur Berechnung komplexer Modelle beträchtlich abnehmen [9,57-59]. Führt die Wahl der Modellierungstechnik zu einer Modifikation des Modellkonzeptes, sollte man gewarnt sein, denn dann ist die gewählte Modellierungstechnik höchstwahrscheinlich falsch.

Die Darstellung in den nachfolgenden Absätzen zu Modellierungstechniken ist keineswegs erschöpfend, aber sie beinhaltet die in der Gesundheitsökonomie am meisten verwendeten

Techniken. Als weiterführende Literatur werden die Leser auf Brennan's Taxonomie der Modellstrukturen [6], Koopman's Untersuchung von Infektionserkrankungsmodellen [60] und Stahl's Übersicht zu Modellierungsmethoden [56] verwiesen.

10.1 Aspekte der Wahl der Modellierungstechnik

Obgleich viele Modellcharakteristika bei der Wahl der Modellierungstechnik berücksichtigt werden könnten [6], sind zwei in den gesundheitsökonomischen Evaluationen besonders prominent. Der erste Aspekt ist die Ebene auf welcher die Population modelliert wird (die gesamte Gruppe oder die einzelnen Individuen der Gruppe). Der zweite Aspekt ist das Basiskonzept welches benutzt wird, um das System zu repräsentieren (basierend auf den Zuständen in welchem sich die Menschen befinden können oder basierend auf den Ereignissen, die ihnen widerfahren können).

10.1.1 Kohortenmodelle versus Individualmodelle

Allgemein gesprochen können Modelle in zwei Kategorien aufgeteilt werden. Die einen simulieren die individuellen Patienten, die anderen unterscheiden nicht zwischen Individuen und verwenden stattdessen den Kohortenansatz. In Kohortenmodelle werden die Individuen zu einer Gruppe aggregiert, welche somit zur Analyseeinheit wird. Im Verlauf der Zeit bricht diese Gruppe entsprechend den modellierten Ereignissen in vordefinierte Untergruppen auf. Die Initialgruppe ist durch einen Satz von Charakteristika definiert (z. B. Alter, Geschlecht, Vorhandensein einer Erkrankung). Die nachfolgenden Subgruppen sind ebenso durch eigene Charakteristika definiert (z. B. neues Krankheitsstadium). Individualmodelle dagegen betrachten die Ereignisse eines jeden Patienten separat, auch wenn sie die Ergebnisse auf dem Populationsniveau darstellen. Jedes Individuum hat eine einzigartige Zusammenstellung von Charakteristiken und wird unter Erfassung seiner persönlichen Ereignisse und Erfahrungen über den Zeitverlauf modelliert. Obgleich diese individuellen Modelle typischerweise mehr Daten, Rechnerkapazität und Laufzeiten [57,58] benötigen, sind sie nicht notwendigerweise schwieriger zu spezifizieren und zu verstehen.

Kohortenmodelle sind die heutzutage in der gesundheitsökonomischen Evaluation am meisten verwendeten Techniken [12,61]. Individuelle Simulation (häufig auch als Mikrosimulation bezeichnet), die früher in der gesundheitsökonomischen Evaluation kaum bekannt war, hat in den zurückliegenden Jahren beträchtlich an Bedeutung gewonnen [57,58]. Trotz weiterhin

bestehender Widerstände gegen ihre Verwendung [62] wird doch weithin akzeptiert, dass eine Simulation auf dem Individuallevel in vielen Situationen mehr Flexibilität und Genauigkeit bieten kann und Mikrosimulationsmodelle für ein breiteres Spektrum an gesundheitswissenschaftlichen Fragestellungen benutzt werden können. Andererseits, wenn die individuellen Patientencharakteristiken und ihre Heterogenität in der Zielpopulation keine große Relevanz für das Entscheidungsproblem haben, kann die Kohortensimulation ausreichend sein, um die wissenschaftliche Fragestellung zu beantworten [3,6,9,55,57,58,63-65].

Für die gesundheitsökonomischen Evaluationen im Auftrag des IQWiG muss die Auswahl der Kohortensimulation gegenüber der Individualsimulation sorgfältig begründet sein. Wenn eine Kohortentechnik angewandt wird, muss gezeigt werden, dass durch die Auswahl des Kohortenansatzes das Modellkonzept nicht in unangemessener Weise verändert wurde und keine Verzerrungen entstehen, die wesentliche Abweichungen von der Realität darstellen. Wenn eine Simulation auf Individualniveau eingesetzt wird, muss dargelegt werden, dass geeignete Verteilungen und Kovarianzmatrizen für die Modellparameter genutzt wurden und dass die Annahmen bezüglich dieser Verteilungen keine unrealistischen Vereinfachungen der Wirklichkeit darstellen.

10.1.2 Ereignisbasierte versus zustandsübergangsbasierte Simulation

Allgemein gesprochen, fanden bislang zwei Ansätze Anwendung, um das untersuchte gesundheitsökonomische System zu konzeptionalisieren. Die für Modellierungen im Allgemeinen am häufigsten verwendete Technik betrachtet die Ereignisse, die eintreten können – Diskrete Ereignissimulation [66] und verwandte Techniken [67]. Im Rahmen gesundheitsökonomischer Evaluationen ist jedoch heutzutage das sogenannte Zustandsübergangskonzept (Markov-Modellierung) weitaus stärker verbreitet [61]. Zustandsübergangsmodelle wurden in der gesundheitsökonomischen Evaluation in den frühen 1980er-Jahren eingeführt [68] und blicken auf eine lange Geschichte im Bereich der Entscheidungsfindung im Gesundheitswesen zurück, einschließlich klinischer und ökonomischer Anwendungen [69].

In einem Markov-Modell durchlaufen die Patienten unterschiedliche, sich gegenseitig ausschließende Gesundheitszustände (Markov-Zustände) in Zeitintervallen (Markov-Zyklen).

Während dieses Prozesses werden kumulative Ereignisse, Gesundheitseffekte (z. B. Lebensjahre, QALYs) und Kosten berechnet und können somit zwischen den interessierenden Interventionen verglichen werden [70]. In einem Zustandsübergangsmodell startet die gesamte Kohorte zum Zeitpunkt null entweder vom gleichen Initialzustand oder entsprechend einer vorgegebenen Initialverteilung von verschiedenen Markov-Zuständen. In jedem Modellzyklus wird die Verteilung der Patienten auf die Gesundheitszustände des Modells unter Berücksichtigung der geeigneten Übergangswahrscheinlichkeiten neu berechnet. Indem diese Analyse über viele Zyklen durchgeführt wird, entsteht ein Profil dazu, wie viele Patienten sich im Laufe der Zeit in dem jeweiligen Zustand befinden.

Zustandsübergangsmodelle basieren auf mehreren vereinfachenden Annahmen und haben Limitationen. Patienten (entweder als Teile einer Kohorte oder als Individuen) werden über die Zustände beschrieben, in denen sie sich befinden können [69]. Diese diskreten Zustände schließen sich gegenseitig aus und werden als homogen angenommen (z. B. gilt jede Person in einem gegebenen Zustand als identisch). Daraus resultiert, dass eine ausreichend große Zahl von Zuständen gewählt werden muss, um das vollständige Spektrum einer Erkrankung darzustellen. Wenn eine Person zwei oder mehr Eigenschaften besitzt (z. B. früherer Myokardinfarkt und Schlaganfall) müssen Markov-Zustände für solche kombinierten Zustände geschaffen werden. Darüber hinaus bewegt sich die Zeit in gleichlangen diskreten Sprüngen, sodass eine ausreichend kleine Zykluslänge zu wählen ist um auch mehrfache Zustandswechsel innerhalb kurzer Zeit zu ermöglichen, sofern diese in der Realität vorkommen.

Eine wichtige Annahme bei Markov-Modellen ist, dass es kein Gedächtnis gibt (die Markov-Annahme). Mit anderen Worten, Personen können sich von einem Zustand in den anderen bewegen, wobei sie alle Charakteristika des neuen Zustandes erwerben und alle Charakteristika zurückliegender Zustände verlieren. Die Wahrscheinlichkeit des Wechsels vom einen Zustand in den anderen hängt einzig und allein vom gegenwärtigen Gesundheitszustand ab. Insbesondere ist diese Wahrscheinlichkeit weder abhängig von vorhergehenden Erfahrungen noch von der Aufenthaltsdauer in einem bestimmten Zustand. Das bedeutet, dass relevante Erfahrungen wie vorhergehende Ereignisse in die Zustandsbeschreibungen einbezogen werden müssen. Beispielsweise müssen sogenannte „Tunnelzustände“ erstellt werden, um die Zeit, die ein Patient bereits in einem spezifischen

Krankheitsstadium zugebracht hat, abzubilden. Dieser Umstand kann zu einer großen Anzahl von Zuständen führen und das Markov-Modell unhandlich machen. In derartigen Fällen kann eine Simulation auf der Individualebene mittels Monte-Carlo-Simulation erster Ordnung sinnvoll sein. Bei dieser Art der Evaluation durchläuft anstelle einer Kohorte eine große Anzahl einzelner Patienten das Modell. Dies ermöglicht die Nachverfolgung der Geschichte eines jeden einzelnen Individuums während der Simulation und überwindet damit die restriktive (und häufig unrealistische) Markov-Annahme. In einem derartigen Modell können weitere Übergangswahrscheinlichkeiten, Kosten und sonstige Parameter als Funktion der gesamten Krankengeschichte modelliert werden [69].

Die ereignisbasierte Simulation [66] konzeptionalisiert die Realität im Sinne der Ereignisse, die auftreten können. Dies erweist sich als flexibles Konzept, das eine natürliche Abbildung des Krankheitsverlaufes, der Behandlungsmuster, der Compliance und anderer relevanter Faktoren erlaubt. Da die Zeit explizit und kontinuierlich berücksichtigt wird und sämtliche in der Realität möglichen Ereignisse ohne *A-priori*-Restriktionen in beliebiger Reihenfolge modelliert werden können, ist es möglich, die Erfahrung der Patienten sehr genau zu erfassen. Ereignisse können somit Veränderungen der klinischen Eigenschaften, medizinische Prozeduren und Arztbesuche, die Inanspruchnahme anderer Ressourcen, das Patientenverhalten, Entscheidungspunkte und sogar Veränderungen des Systems im Ganzen (z. B. die Einführung einer neuen Behandlung oder die Veränderungen der Erstattung) wiedergeben. Darüber hinaus erlauben diese Simulationen den Individuen nicht nur ihre eigenen Charakteristika, sondern auch ihre eigene persönliche Krankheitsgeschichte zu „tragen“. Dadurch kann im Modell verfolgt werden, was bereits geschah, und diese Information kann dazu verwendet werden, den nachfolgenden Ereignisverlauf zu modifizieren. Somit können die Zeit seit Beginn eines Zustandes, die Länge der Behandlung, das Ansprechen auf eine Behandlung, vorhergehende Ereignisse und Entscheidungen sowie andere relevante zurückliegende Geschehnisse den nachfolgenden Verlauf bestimmen, wenn dies gewünscht wird. Da Ressourcen und ihre Eigenschaften, einschließlich Kapazitäten, explizit modelliert werden, stellen diese Simulationen eine sehr realistische Plattform für die Abschätzung der Kosten dar. Es kann festgestellt werden, dass ereignisbasierte Modelle sehr leistungsstarke Instrumente darstellen, die gegenwärtig in der Gesundheitsökonomie zu wenig Anwendung finden.

Ausgehend von der Flexibilität und der Realitätsnähe ereignisbasierter Simulationen wird empfohlen, dass diese Modelle in derselben Weise für IQWiG-Bewertungen genutzt werden sollten wie zustandsübergangsbasierte Modelle. Wenn beispielsweise ein Individuum innerhalb eines einzigen Tages ein Schlaganfallereignis, eine Krankenhauseinweisung und den Tod erleidet und dies durch ein Modell erfasst werden soll, könnte ein ereignisbasierter Ansatz die beste Wahl sein, da dieser eine genauere Zuordnung von Kosten und Diskontierungen erlaubt [58].

Es muss eine detaillierte Begründung für die Wahl zwischen einem Zustandsübergangs- und einem ereignisbasierten Modellierungsansatz erfolgen und es muss dargelegt werden, dass die Wahl einer bestimmten Modellierungstechnik zu keinen unangemessenen Veränderungen des Modellkonzepts oder unakzeptablen Realitätsverzerrungen führt.

10.2 Techniken

10.2.1 Entscheidungsbaumanalysen

Ein Entscheidungsbaum ist eine visuelle Darstellung aller möglichen Optionen und Konsequenzen, die sich als Folge einer untersuchten Technologie ergeben können [71]. Jeder der untersuchten Interventionen folgen Verzweigungen, die die möglichen Ereignisse mit den dazugehörigen Eintrittswahrscheinlichkeiten darstellen. Hierbei können die Wahrscheinlichkeiten nicht nur von den unterschiedlichen Strategien, sondern auch von Patientencharakteristika abhängen (z. B. Subgruppen mit unterschiedlichen Risikofaktorprofilen). Am Ende des Entscheidungsbaums führt jeder Pfad zu einem Outcome, z. B. dem Endpunkt einer klinischen Studie, einem kombinierten Score für Symptome und Nebenwirkungen, Lebensqualität, krankheitsfreiem Überleben oder Überleben schlechthin. Für jede der alternativen Interventionen kann der Erwartungswert der klinischen und ökonomischen Konsequenzen als ein gewichteter Mittelwert aller möglichen Konsequenzen berechnet werden, wobei die Pfadwahrscheinlichkeiten als Gewichte verwendet werden. Entscheidungsbaumanalysen eignen sich sehr gut für sehr einfache Entscheidungsprobleme, bei denen Ereignisse mit limitierten Rekursionen und kurzem fixen Zeithorizont analysiert werden. Entscheidungsbäume können entweder mittels Kohortensimulation oder Simulation auf Individualebene evaluiert werden.

10.2.2 Diskrete Ereignissimulation

Auf dem Gebiet der Gesundheitsökonomie hat der Einsatz der diskreten Ereignissimulation in den letzten Jahren an Bedeutung gewonnen [56]. Nichtsdestotrotz befassen sich existierende Bücher [72] und Manuale [73,74] mit Beispielen aus vielen anderen Gebieten, werden jedoch – abgesehen von der Betrachtung der Organisation von Gesundheitsdiensten – nicht auch in der Gesundheitsökonomie eingesetzt [75-79]. Allerdings sind die Grundlagen der diskreten Ereignissimulation in der nichtmedizinischen Literatur sehr gut dokumentiert [80]. In diesem Abschnitt werden die wichtigsten Elemente dieser Technik zusammen mit ihrer Anwendung im Rahmen gesundheitsökonomischer Evaluationen kurz dargestellt.

Die diskrete Ereignissimulation operiert auf der Ebene individueller Entitäten; diese stellen die Objekte dar, welche die im Modell berücksichtigten Ereignisse erfahren. In einer diskreten Ereignissimulation kann es sehr viele Arten von solchen Objekten geben und die Entitäten können sogar während der Simulation ihren Typ verändern. In gesundheitsökonomischen Modellen stellt der Patient eine nahezu universelle Entität dar, aber in einem tiefer gehenden Modell könnte die Entität auch auf der Ebene der Körperorgane definiert sein. Andere bekannte Entitäten sind z. B. Angehörige der Heil- und Pflegeberufe und sogar leblose Objekte. Entitäten werden zu jedem Zeitpunkt der Simulation durch ein sogenanntes „Enter-Ereignis“ gebildet und einem Typ zugeordnet. In gesundheitsökonomischen Evaluationen ist es üblich, sämtliche Entitäten bereits zu Beginn der Simulation zu kreieren, aber dies ist nicht unbedingt erforderlich. Entitäten können aufgrund von Neuerkrankung, Geburt, Einwanderung oder anderen ähnlichen Mechanismen auch später in die Simulation eintreten. Die Entitäten verbleiben in der Simulation (auch wenn sich bei ihnen nichts ereignet) bis ein sogenanntes „End-Ereignis“ auftritt, bei dem das finale Processing durchgeführt wird und sie die Simulation verlassen. Obgleich der Tod die häufigste Art eines solchen End-Ereignisses ist, können die Entitäten auch aus anderen Gründen die Simulation verlassen, z. B. aufgrund des Endes der zeitlichen Betrachtung, des Verlustes der Indikation, durch Emigration usw.

Entitäten zeichnen sich dadurch aus, dass ihren Attributen spezielle Werte zugeordnet werden. Alle Entitäten teilen einen gemeinsamen Satz von Attributen (z. B. würden alle das Attribut Alter haben, wenn dies ein Teil des Modells wäre); den individuellen Entitäten werden aber ihre eigenen eindeutigen Werte für jedes Attribut zugeordnet. Diese individuellen Charakteristika werden im Verlauf der Simulation aktualisiert. Veränderungen ergeben sich

aus dem natürlichen Zeitablauf oder aufgrund von Ereignissen, die die Entitäten erleben (z. B. medizinische Komplikationen, Arztbesuche, Änderungen in der Befolgung des verordneten Handlungsregimes), oder aufgrund von anderen Effekten, die das gesamte System betreffen (z. B. Veränderungen in der Erstattung, Einführung eines neuen Prozesses), bzw. aufgrund von Geschehnissen, die andere Entitäten betreffen (z. B. Tod einer Pflegekraft). Aspekte der Krankheitsgeschichte eines jeden Individuums können als Attribute gespeichert werden, wenn sie das Auftreten der nachfolgenden Ereignisse und Versorgungsentscheidungen oder andere Attribute und Modellierungsaspekte beeinflussen. Das Modell kann auf diese Attributwerte zugreifen und sie können zum Zweck der Ergebnispräsentation oder für weiterführende Analysen ausgegeben werden. Dadurch kann Heterogenität in der Population, den Behandlungseffekten und den Outcomes einfach in die Simulation eingebracht werden können, da es sich hierbei um individuelle Attribute handelt.

In einer diskreten Ereignissimulation wird die Zeit explizit modelliert und verläuft wie in der Realität kontinuierlich. Dies bedeutet, das zeitliche Auftreten der Ereignisse unterliegt keinerlei Beschränkungen – Ereignisse können sogar simultan auftreten (z. B. an ein und demselben Tag) und konkurrierende Risiken können angemessen berücksichtigt werden. Es besteht keine Notwendigkeit für sogenannte Halbzyklus-Korrekturen oder die Anwendung anderer Maßnahmen zur Reduktion der durch die Vorgabe fester Zeitzyklen bedingten Abweichungen. Dem Analytiker wird ermöglicht, alle bestehenden Zeitabhängigkeiten sachgemäß zu implementieren und das zeitliche Auftreten der Ereignisse sowie die Dauer der Krankheitszustände etc. zu verfolgen. Verzögerungen — sowohl geplante (z. B. einen Tag in ein Krankenhaus wegen eines Eingriffs) als auch fremdbestimmte (z. B. Warten auf ein verfügbares Bett) — können problemlos integriert werden. Im Rahmen ökonomischer Modelle ermöglicht die explizite Handhabung der Zeit eine vollständige, kontinuierliche Diskontierung.

Eine besondere Stärke der diskreten Ereignissimulation ist, dass Ressourcen ebenfalls explizit modelliert werden. Sie werden als besondere Ereignisse angesehen, bei denen eine Ressource in Anspruch genommen wird. Diese Ressourcen können sowohl Personen (z. B. Ärzte, Krankenschwestern, Therapeuten), Güter (Medikamente, Katheter), Organe (z. B. Lebertransplantate), Platz (z. B. Krankenhausbetten) oder beliebige andere Dinge sein, die als Ressource verstanden werden. Den Ressourcen sind Kosten pro Einheit zugeordnet, welche

durch den Gebrauch pro Stunde usw. spezifizierbar sind und die zeitabhängig sein können (z. B. teurer außerhalb der normalen Arbeitszeit); selbst die Kosten von Inaktivitätszeiten können spezifiziert werden. Ressourcen beschreiben auch Kapazitätseinheiten (z. B. Betten in einer Intensivstation oder Ärzte innerhalb einer Region). So kann es geschehen, dass ein bestimmter Typ einer Ressource zum Engpass wird und Warteschlangen entstehen, weil Individuen auf das Freiwerden der benötigten Ressource warten müssen. Sollten solche Ressourcenbeschränkungen ein relevantes Merkmal eines Entscheidungsproblems in einer IQWiG-Evaluation darstellen, wird empfohlen, die diskrete Ereignissimulation anstatt eines Markov-Modells einzusetzen.

Ein anderes Merkmal der diskreten Ereignissimulation ist es, dass Individuen miteinander interagieren können [3,6,74]. Dies kann verschiedene Formen annehmen. Beispielsweise erlaubt die diskrete Ereignissimulation die Spezifizierung von eingeschränkten Ressourcen, bei der individuelle Entitäten um diese Ressourcen konkurrieren und Warteschlangen bilden, wenn die benötigte Ressource nicht frei ist. Im Gesundheitswesen, in dem eingeschränkte Ressourcen allgegenwärtig sind, kann die Bewertung, in welcher Weise neue Interventionen entlastend auf limitierte Ressourcen wirken, bedeutsam sein. Andere Beispiele beinhalten Verhaltensänderungen. Beispielsweise kann es im Zuge der zunehmenden Erfahrung der Mediziner mit einer Gesundheitstechnologie durch die Behandlung vieler Patienten zu einer Änderung der Effizienz und Behandlungsmuster bei zukünftigen Patienten kommen. Schließlich ist es für Modellierungen auf das Gebiet der infektiösen Erkrankungen, bei denen die Übertragung von Infektionen eine unmittelbare Funktion des Infektionsstatus der Individuen in der Gemeinschaft darstellt, äußerst sachdienlich, eine Möglichkeit zur Spezifizierung der Interaktionen und Ansteckungswege zwischen den Entitäten zu haben (siehe Kapitel 10.2.5).

Diskrete Ereignissimulationen und auf der Individualebene analysierte Markov-Modelle benutzen verschiedenartigste Gleichungen um den Modellierungsprozess zu kontrollieren. Dies können lineare Funktionen, logistische Regressionsfunktionen, Hazardfunktionen, und Cox-proportional-Hazardfunktionen oder andere Funktionsarten sein, die mathematisch erfasst werden können. Eine diskrete Ereignissimulation verwendet auch Verteilungen wie Exponential-, Weibull-, Beta-, Gammaverteilungen usw., um die Variabilität der Inputparameter und Ereigniszeiten etc. abzubilden. Zusammen erzeugen diese das

stochastische Verhalten des Modells. Zufallszahlen müssen mittels hochentwickelter Algorithmen erzeugt werden, die eine ausreichende Zufälligkeit in der Sequenz garantieren

Bedingt durch die stochastische Natur der diskreten Ereignissimulation können einzelne Durchläufe des Modells sehr untypische Ergebnisse liefern. Deshalb ist es erforderlich, die Durchläufe viele Male zu replizieren. Jede Replikation verwendet einen unterschiedlichen Satz von Zufallszahlen und somit sind alle Prozesse, die Zufallszahlen verwenden, einschließlich der Generierung der Population und aller ihrer Erfahrungen, einzigartig. Dies ergibt einen unterschiedlichen Satz von Resultaten für jede einzelne Replikation. Nach Durchführung vieler derartiger Replikationen werden sich die Ergebniswerte ihrem Mittelwert annähern und ergeben so eine stabilere Schätzung der Resultate. Eine häufig gestellte Frage lautet „Wie viele Replikationen sind notwendig?“. Auf diese Frage gibt es keine allgemeingültige Antwort, weil dies von der Variabilität innerhalb des Modells und zwischen den Patienten etc. abhängig ist. Der beste Weg, dies zu beurteilen, ist ein empirischer: die Untersuchung der Standardabweichungen der kritischen Ergebnisse zwischen den Replikationen. Wenn die Größe der Standardabweichung akzeptabel ist, war die Anzahl der Replikationen ausreichend.

Während es möglich ist, eine diskrete Ereignissimulation so zu programmieren, dass sie sich wie ein Kohortenmodell verhält – was eine Versuchung darstellt, wenn nicht genügend Daten zur Verfügung stehen –, erlaubt die Flexibilität dieser Technik die simultane Spezifikation multipler Komplexitätsebenen in ein und demselben Modell. Im Falle stark limitierter Daten, kann ein Modell mit einem einfachen Satz von Inputparametern versehen werden (z. B. generelle Verwendung des durchschnittlichen Behandlungseffekts anstatt aus der Verteilung der Behandlungseffekte gezogenen Werten), dasselbe Modell kann aber ebenso für die Analyse anspruchsvollerer Szenarien verwendet werden. Dies ermöglicht dem Analytiker eine Aussage über die potenzielle Veränderung der Ergebnisse, falls ausführlichere Daten zur Bestückung des Modells zur Verfügung stünden (z. B. was würde passieren, wenn der Behandlungseffekt auf der Ebene von Individuen variieren kann).

Die Anwendung der diskreten Ereignissimulation und deren Gebrauch in der Gesundheitsökonomie sind verschiedentlich beschrieben worden [3,54,56,58,64,81]. Da die diskrete Ereignissimulation den Rahmen für eine sehr natürliche Darstellung der Krankheit und des Ressourcenverbrauchs bietet, ist sie gut dafür geeignet, die Folgekosten von

Gesundheitsinterventionen zu modellieren. Darüber hinaus ermöglicht moderne Software eine sehr transparente Darstellung der Beziehungen innerhalb des Modells, z. B. der Patientenströme innerhalb der Simulation. Softwareanwendungen wie Arena® [74] erlauben die Programmierung der diskreten Ereignissimulation mithilfe von Flowchart-Methoden, die denjenigen gleichen, mit denen die Kliniker bereits vertraut sind. Diese ermöglichen eine visuelle Darstellung der Patientenströme und möglichen Outcomes. Da die diskrete Ereignissimulation ein breites Spektrum von Zwischen- und Endergebnissen erzeugt, ist es möglich, Outcomes auf verschiedenen Ebenen zu bestimmen. Dies ist für Validierungszwecke nützlich und ermöglicht einen umfassenderen Ergebnissatz, der sich auf die Belange verschiedener Entscheidungsträger und Analytiker richtet.

10.2.3 Markov-Modelle

Markov-Modelle sind mächtige und einfach anwendbare Instrumente zur Modellierung von Prävention, Diagnose und Behandlung chronischer Erkrankungen, bei denen (1) die Parameter zeitabhängig sind (z. B. Restenose nach einem koronaren Eingriff, Rezidiv nach einer Krebstherapie, AIDS nach dem Beginn mit HAART), (2) der Ereigniszeitpunkt eine Rolle spielt (z. B. Zeit bis zu einem Ereignis wie einem Schlaganfall, Krebsrezidiv etc.) oder (3) Ereignisse wiederholt auftreten können (z. B. Restenose, späte Komplikationen, zweiter Myokardinfarkt) [82].

Es existieren verschiedene gute Anleitungen zur Konstruktion von Markov-Modellen [61,69,83] und der Analyst sei in im Hinblick auf Einzelheiten auf diese verwiesen. Wie oben dargestellt bilden Markov-Modelle die Wirklichkeit in Form einer Abfolge einzelner *Zustände* ab, welche die relevanten gesundheitlichen Zustände des Patienten (z. B. keine Krankheit, Krankheit, Tod) widerspiegeln. In einem Markov-Kohortenmodell [83] wird die Population in die bestehenden Zustände aufgeteilt; anschließend wird anhand des Modells der Anteil der Kohorte, der innerhalb eines festen Zeitintervalls (*Zyklus*) von einem Zustand in den anderen wechselt (*Übergänge*), berechnet. (Obwohl die Beschränkung auf ein festes Zeitintervall nicht zwingend ist [84], werden Markov-Modelle bei Evaluationen im Gesundheitswesen weitgehend auf diese Weise durchgeführt.)

Zustände in einem Markov-Modell schließen sich gegenseitig aus (kein Teil der Population kann in demselben Zyklus zwei Zuständen angehören) und sind vollständig (die Summe der

Kohortenanteile in allen Zuständen zusammengenommen muss in jedem Zyklus 100 % entsprechen). Aus *absorbierenden Zuständen* gibt es keinen Austritt: Der Teil der Kohorte, der diesen Zustand erreicht, verbleibt dort für den Rest der Modellierungsdauer. Tod ist der am häufigsten verwendete absorbierende Zustand. Jeder Zustand wird entsprechend der in der Analyse verwendeten Einheit bewertet (zum Beispiel „Kosten“ in einem ökonomischen Modell).

In Kohortenmodellen wird die Population in jedem Zustand durch einen Satz gemeinsamer Merkmale definiert, welche diesen Teil der Population beschreiben (z. B. Alter, Geschlecht, Krankheitsschwere) und welche sich zwischen den Individuen innerhalb eines Zustands nicht unterscheiden [58,61]. Diese Merkmale können die Übergangswahrscheinlichkeiten von einem Zustand in einen anderen Zustand im nächsten Zyklus beeinflussen (z. B. ist der Anteil der erkrankten Kohorte, der sterben wird, abhängig vom Alter der Kohorte) und sie können Auswirkungen auf die Bewertung des Zustandes haben (z. B. hängen die kumulierten Kosten während eines Zyklus von der Krankheitsschwere ab) [83]. Die Charakteristika jedes Zustands können sich auch im Laufe der Zeit ändern. Markov-Modelle sind für Wissenschaftler im Gesundheitswesen und Ärzte aufgrund des Denkens in Zuständen und Übergangskurven besonders attraktiv. Insbesondere Ärzte sind es gewohnt, in Krankheitsstadien (NYHA-Stadien I-IV bei Herzinsuffizienz, Karzinomstadien I-IV etc.) und in Übergängen mit jährlichen Wahrscheinlichkeiten oder Kaplan-Meier-Überlebenskurven zu denken. Daten werden häufig so dargestellt, dass sie direkt für Markov-Modelle genutzt werden können. Allerdings sollte, wie oben erwähnt, das Format der verfügbaren Daten den Analysten nicht dazu bewegen, ein Markov-Modell gegenüber einem anderen Modelltyp vorzuziehen.

In Kohortenmodellen kann es vorkommen, dass die Charakteristika der Zustände nicht alle wichtigen Eigenschaften enthalten, die für eine spezifische Modellierung der Übergänge erforderlich sind. Dies kann zu erheblichen Verzerrungen führen (z. B. in Einjahreszyklen wird das mittlere Alter üblicherweise um ein Jahr erhöht, obwohl die älteren Personen aufgrund von Krankheit oder Tod ausscheiden können). Dieser Aspekt von Markov-Modellen kann überwunden werden, wenn man diese Heterogenität durch die Einführung zusätzlicher hinreichend homogener Gesundheitszustände berücksichtigt. Bei multiplen Patientencharakteristika kann dies allerdings schnell zu einem expandierenden Satz an

Zuständen führen bis hin zu einer nicht mehr möglichen Implementierung. Dieses Problem kann durch eine Simulation von Individuen im Markov-Modell behoben werden (siehe unten).

Die einfachsten Markov-Modelle sind *Ketten*, welche der Annahme unterliegen, dass die Übergangswahrscheinlichkeiten über die gesamte Modellierungsdauer hinweg konstant sind [69]. Markov-Ketten sind eine angenehme Art zu modellieren, da sie numerisch berechnet werden können [85], indem die zugehörigen Matrizen invertiert werden (d. h., es ist nicht einmal notwendig, das Modell laufen zu lassen). Die Eigenschaft konstanter Übergangswahrscheinlichkeiten ist allerdings im Gesundheitswesen meistens unrealistisch. Zum Beispiel ändert sich die Hintergrundmortalitätsrate von Individuen mit zunehmendem Alter. Da Tabellenkalkulationsprogramme und, noch aktueller, spezielle entscheidungsanalytische Software iterative Berechnungen in Markov-Modellen erlauben, welche nicht der technischen Komplexität der Matrix-Algebra unterliegen, werden in gesundheitsökonomischen Evaluationen überwiegend *Markov-Prozesse* verwendet, welche die Änderung der Übergangswahrscheinlichkeiten im zeitlichen Verlauf erlauben (z. B. das Sterberisiko mit zunehmender Zeit im Modell zu erhöhen) [68].

Ein weiterer Aspekt von Markov-Modellen in Bezug auf Übergangswahrscheinlichkeiten ist die Markov-Annahme [57,61,83], welche besagt, dass die Wahrscheinlichkeiten des Wechsels von einem in den anderen Zustand einzig vom aktuellen Zustand abhängt und nicht von vorhergehenden Zuständen, der Aufenthaltsdauer in Zuständen oder anderen Ereignissen, welche die Population in diesem Zustand in der Vergangenheit erfahren hat. Diese Annahme wird häufig als *Gedächtnislosigkeit* bezeichnet: Das heißt, das Modell hat keine Erinnerung an das, was zuvor mit dem Patienten geschah. Daher muss eine ausreichend große Menge an Zuständen gewählt werden, welche die wesentliche Patientengeschichte umfasst, um ein realistisches Modell der Krankheit zu erstellen.

In Markov-Modellen wird die Zeit durch feste Zeitintervalle spezifiziert, über welche die Übergangswahrscheinlichkeiten angewandt werden. Pro Zyklus findet nur ein Übergang statt. Dies erfordert zusätzliche Transformationen der Wahrscheinlichkeiten, um konkurrierende Risiken korrekt zu behandeln [58]. Der Analyst muss eine ausreichend kleine Zykluslänge wählen, um den kontinuierlichen Zeitfluss in der Wirklichkeit nachzuahmen.

In einem reinen Markov-Modell werden Interaktionen zwischen Individuen nicht betrachtet [3,6,54]. Allerdings ist diese Beschränkung in entscheidungsanalytischen Softwareprogrammen inzwischen aufgehoben und „externe“ Parameter, wie Krankheitsprävalenz, können verwendet und im Modell kontinuierlich aktualisiert werden. Diese Erweiterung kann dazu verwendet werden, Transmissionsmechanismen in die Struktur eines Markov-Modells einzubinden.

Markov-Modelle mit Simulation auf Individualebene werden aufgrund der verwendeten Stichprobenziehung aus Wahrscheinlichkeitsverteilungen oft als *Monte-Carlo-Simulationen* bezeichnet. Sie umgehen viele der Limitationen, die den Kohortenmodellen eigen sind [6,6,61,69]. In diesen Modellen bekommt jedes Individuum seine eigenen Charakteristika zugewiesen und so kann die Historie eines jeden Individuums gespeichert werden. Damit kann der Heterogenität in Behandlungseffekten und Outcomes Rechnung getragen werden. Im Gegensatz zu Kohortenmodellen sind Monte-Carlo-Simulationen stochastischer Natur, und alle Elemente der Simulation können durch Wahrscheinlichkeitsverteilungen bestimmt werden. Der Patientenfluss durch das Modell wird durch Stichprobenentnahmen dieser Wahrscheinlichkeitsverteilungen bestimmt. Eine Schwierigkeit tritt auf, wenn diese Modelle ein Individuum zu einem Zeitpunkt durchrechnen und dabei alle Übergangswahrscheinlichkeiten jedes Zyklus anwenden, obwohl das Individuum in den meisten Zyklen keine Übergänge aufweist. Dies verlangsamt den Berechnungsprozess erheblich.

Ähnlich wie diskrete Ereignissimulationen können Markov-Modelle mit Simulation auf Individualebene viele Arten an Gleichungen nutzen, um Modellierungsprozesse zu kontrollieren (z. B. lineare Funktionen, logistische Regressionsfunktion, Hazardfunktionen, Cox-Proportional-Hazards-Funktionen etc.).

Markov-Modelle können überdies in unterschiedlicher Weise modifiziert werden, um die Limitationen eines reinen Markov-Modells zu umgehen. Zum Beispiel werden „Tunnelzustände“, „Halbzyklus-Korrekturen“, multiple Unterzustände, externe Makros und andere Möglichkeiten entscheidungsanalytischer Softwareprogramme routinemäßig in der Praxis genutzt. Allerdings wirken diese Funktionen einem der wichtigsten Vorteile von Markov-Modellen entgegen – ihrer Einfachheit.

Inzwischen hat die Entwicklung von moderner Software dazu geführt, dass die Programmierung von Markov-Modellen sowohl für die Kohortensimulationen als auch die Simulation auf Individualebene vereinfacht wurde und dass Modellierungsprozesse auf intuitive Weise dargestellt werden können, was die Transparenz erleichtert. Darüber hinaus gibt es spezielle Instrumente, welche die Visualisierung der Zustandsverteilungen im Zeitablauf (die sogenannte *Markov-Spur*), Sensitivitätsanalysen, die Darstellung von Unsicherheiten, Schnittstellen zu Tabellenkalkulationsprogrammen und zur Makrosprache unterstützen und viele andere effiziente Instrumente.

10.2.4 Agentenbasierte Simulation

Die Agentenbasierte Simulation ist eine mächtige Weiterentwicklung der diskreten Ereignissimulation, wird aber seltener als Modellierungstechnik im Gesundheitswesen verwendet [53,67,86-88]. Der Hauptunterschied zur Standardform der diskreten Ereignissimulation ist, dass in einem agentenbasierten Modell die Entitäten (die man hier als „Agenten“ bezeichnet) sich gegenseitig und ihre Umgebung wahrnehmen [67]. Diese Technik ist hervorragend für die Modellierung von Infektionskrankheiten geeignet, bei denen die Übertragung von Infektionen wichtig ist. Zu den relevanten Anwendungsmöglichkeiten gehören alle Bereiche, in denen soziales Verhalten eine Schlüsselrolle als Einflussfaktor auf Outcomes darstellt. Obwohl die einfache diskrete Ereignissimulation viele dieser Dinge bewältigen kann, ist die agentenbasierte Simulation ein natürlicherer Ansatz, insbesondere dort, wo geografische oder räumliche Betrachtungen eine Rolle spielen. Sie erleichtert den Programmierungsprozess und ermöglicht eine transparentere Darstellung der Funktionsweise des Modells.

Ein Agent, wie eine Entität, repräsentiert häufig Patienten oder Personen mit einem Risiko für die jeweilige Krankheit. Wie eine Entität können Agenten aber auch so eingesetzt werden, dass sie andere Dinge, wie medizinisches Fachpersonal, Organisationen oder Gemeinschaften, repräsentieren. Agenten können im Ablauf der Simulation ihr Verhalten modifizieren und das Verhalten anderer Agenten beeinflussen. Daher ist eine der entscheidenden Komponenten eines Agenten seine „*capability ... to make independent decisions [and] ... to be active rather than purely passive*“ [67]. Die agentenbasierte Simulation ermöglicht auch die Festlegung räumlicher Beziehungen zwischen den Agenten.

Wie bei der diskreten Ereignissimulation hat neue Software die Programmierung von agentenbasierten Simulationen vereinfacht, die eine unmittelbare visuelle Darstellung der Modellprozesse ermöglicht sowie die Transparenz der Modelle erhöht [53]. Diese visuellen Hilfsmittel sind nicht nur wertvoll für den Entscheidungsträger, sondern auch für den Analysten bei der Fehlersuche im Programm.

Einige Szenarien, bei denen agentenbasierte Simulation in Betracht kommen [53], sind:

- Das Verhalten von Individuen ist nicht linear und wird möglicherweise am besten durch nicht kontinuierliche Funktionen beschrieben (z. B. Schwellenwerte).
- Individuelle zeitliche Korrelationen, Adaptierungen und Lernen sollen dargestellt werden.
- Interaktionen zwischen Individuen sind heterogen.
- Netzwerk-Effekte können vorhanden sein.
- In dem System sind Schwankungen möglich, wobei ein stabiler Zustand, basierend auf durchschnittlichen Verhältnissen, unrealistisch ist.

10.2.5 Transmissionsmodelle

Bei der Modellierung von Infektionskrankheiten in den Bereichen Epidemiologie und Public Health stellen Transmissionsmodelle eine weit verbreitete und anerkannte Technik dar [60]. Transmissionsmodelle [89] können mit verschiedenen Komplexitätsgraden konstruiert werden und können entweder deterministisch oder stochastisch sein [55,60]. In ihrer einfachsten Form basieren deterministische Modelle auf Berechnungen einer Reihe von Differenzialgleichungen, welche die infektiösen Zustände (*Kompartments*) der Population (z. B. suszeptibel, infiziert, immun), den Grad der Ansteckungsgefahr der Krankheit und die unterschiedlichen Muster in der Population beschreiben. Zum Beispiel können die Effekte von Impfungen, welche die Suszeptibilität der Population und eventuell die Stärke der Ansteckungsgefahr sowie andere Faktoren ändern können, dabei sowohl in Bezug auf die direkten Effekte der Impfung (d. h. der Schutz für das geimpfte Individuum) als auch hinsichtlich ihrer indirekten Effekte (d. h. die Herdenimmunität) einbezogen werden. Deterministische Kompartimentmodelle der Krankheitsübertragung, welche im Wesentlichen

Kohortenmodelle sind, benötigen allerdings eine Reihe vereinfachender Annahmen, welche in bestimmten Situationen nicht akzeptabel sind [55,60]. Die einfachsten Modelle setzen eine komplette Uniformität der Population voraus sowie unmittelbaren Kontakt und sofortige Infektion, sofortige und gedächtnislose Vermischung, Uniformität bezüglich der infektiösen Ereignisse und Uniformität bezüglich der Genesung von der Infektion [60]. Diese Annahmen sind häufig nicht haltbar, wenn Entscheidungsträgern Informationen über die ökonomischen Konsequenzen der jeweiligen Interventionen angeboten werden. Einige der Annahmen in Transmissionsmodellen können bei anspruchsvollerer Modellierung der Krankheitsübertragung gelockert werden. Techniken, die speziell für solche Bereiche entwickelt wurden, bieten geeignetere und präzisere Modellierungsansätze [60].

10.3 Empfohlene Technik

Das IQWiG gibt keine Empfehlung für die Verwendung einer bestimmten Modellierungstechnik. Die entscheidende Überlegung bei der Beurteilung der Angemessenheit einer Modellierungstechnik ist, ob diese genaue Informationen zu relevanten Kosten liefern kann. Dies beinhaltet mehr als die einfache Berechnung eines Punktschätzers für die erwarteten ökonomischen Konsequenzen der verschiedenen Gesundheitstechnologien. Benötigt wird auch – so vollständig wie möglich – die wahrscheinliche Spanne von Ergebnissen. Ferner ist die Transparenz beim Modellieren wichtig, allerdings ist der Standpunkt des IQWiG, dass Transparenz nicht Vorrang hat vor einer ausreichend detaillierten Modellierung. Mit angemessener Dokumentation sollte es möglich sein, jede Technik angemessen transparent darzulegen. Es obliegt des Analytikers, die Wahl ihrer Technik und Software zu begründen und die Implikationen dieser Wahl zu bewerten.

Das IQWiG gibt ebenfalls keine Empfehlung dazu ab, welche Software oder Programmiersprache beim Modellieren verwendet werden sollte. Die Beschränkung der Modelle auf solche, die vollständig mit gängigen Tabellenkalkulationsprogrammen berechnet werden können, ist in Hinblick auf den benötigten Informationsgrad und die Bedeutung der evaluierten Forschungsfragen nicht angemessen.

Abschließend sollte festgehalten werden, dass unabhängig von der Wahl der Modellierungstechnik valide Modelle ggf. Vorhersagen darüber treffen können, wie sich die in klinischen Studien ermittelten Nutzeneffekte unter anderen Bedingungen oder

Zeitperspektiven darstellen würden. Modelle können jedoch in keinem Fall neue Nutzeneffekte generieren.

11 Umgang mit Unsicherheit und Variabilität

11.1 Arten

Es existieren zwei hauptsächliche Arten von Variabilität in ökonomischen Modellen [71]. Es gibt eine natürlich vorkommende Variabilität, die von Interesse für die Modellierer ist, und es gibt eine Variabilität, die man gerne vermeiden würde, die sich aber als unumgänglich gestalten kann. Der Umgang mit beiden Arten unterscheidet sich ausgesprochen. Im ersteren Fall ist der Sachverhalt Gegenstand von Studien, während es sich im zweiten Fall um einen Sachverhalt handelt, der, soweit möglich, eliminiert werden muss und dessen verbleibendes Residuum zu quantifizieren ist.

Kosten können unter anderem in Abhängigkeit vom Ort der Leistungserbringung, von den unterschiedlichen Patientenkollektiven und von den jeweiligen Bedingungen, unter welchen sie anfallen, variieren. Somit hängt die natürliche Variabilität von den Charakteristika des Programms, den Patienten, dem Kontext und den Stückkosten ab. Anzumerken sei hier, dass diese Elemente mit Sicherheit bekannt sein können (z. B. Initialdosis) oder aber man sich unsicher über ihren Wert sein kann (z. B. Tageskosten eines Krankenhausaufenthalts). In beiden Fällen sind diese Kostendeterminanten für die Analyse relevant und somit zu eruieren. Dies wird durch Sensitivitätsanalysen unternommen, die univariat (Veränderung jeweils einer einzigen Determinanten) oder multivariat sein können. Die Letztere ist von größtem Nutzen, wenn die Analyse anhand spezifischer, durch einen Satz von Variablenwerten definierter Szenarien erfolgt.

Als Teilschritt der Untersuchung der Variabilität von Interesse sollte der Modellierer die Entscheidungen zur Modellstruktur und zu den durchwegs getroffenen Annahmen berücksichtigen. Diese können unter anderem in Zusammenhang mit der Konzeptualisierung der gesundheitlichen Intervention, den einbezogenen klinischen Ereignissen, dem für die Modellierung verwendeten Zeithorizont, den funktionalen Beziehungen, der Auswahl der berücksichtigten Kosten und Nutzen sowie den Methoden zur Messung und Bewertung der Outcomes stehen [34,90,91]. Diese Elemente können dem Modellierer invariant erscheinen, jedoch besteht üblicherweise in einigen Bereichen Uneinigkeit hinsichtlich des angemessenen Ansatzes [92] (z. B. Diskontierungsrate). Falls andere plausible Annahmen für die wichtigsten

Komponenten des Modells gemacht werden könnten, sollte die resultierende Variabilität einer formalen Untersuchung unterzogen werden.

Variation kann auch aus der statistischen Analyse des Datenmaterials (d. h. Schätzungen aus Stichproben mit endlicher Stichprobengröße, die notwendigerweise nie frei von Unsicherheit sind), aus der Modellierungstechnik oder aus individueller Variation resultieren. So können Kosten variieren, weil Unsicherheit über einige ihrer Determinanten besteht, wie beispielsweise die über die Zeit hinweg eingenommene Dosierung (hervorzuheben ist, dass dasselbe Element mit beiden Arten von Variabilität behaftet sein kann), oder weil das Modell stochastisch ausgelegt ist (es verwendet Zufallszahlen bei den Monte-Carlo-Ziehungen). Diese Art von Variabilität ist nicht an sich von Interesse und deshalb werden verschiedene Techniken angewendet, um einen größtmöglichen Anteil davon zu eliminieren. Diese Techniken zur Varianzreduktion (z. B. Klonen) werden in der Literatur ausführlich beschrieben [59,93,94]. Eine verbreitete Technik ist die Vergrößerung der Stichprobengröße durch Simulation zusätzlicher Patienten oder durch mehrmalige Modelldurchläufe. Die nach Anwendung all dieser Techniken verbleibende Varianz muss in allen Fällen quantifiziert werden. Dies kann mittels Anwendung unterschiedlicher statistischer Methoden auf die Modellergebnisse erfolgen oder durch Analysetechniken wie die probabilistische Sensitivitätsanalyse.

11.2 Quantifizierung von Unsicherheit

Die in der Evaluation von Gesundheitsleistungen angewendeten Modelle benötigen meistens eine Datenzusammenführung aus verschiedenen Quellen, wie beispielsweise klinischen Studien, Beobachtungsstudien, Datenbankauswertungen, Befragungen usw. Somit ergibt sich in allen Modellen ein gewisser Grad an Unsicherheit [90]. Die Ursachen dieser Unsicherheit lassen sich auf mehrere Arten kategorisieren [91,95-97].

Die Unsicherheit der Parameter stammt daher, dass der Wert einiger Modellparameter aus Stichproben mit endlicher Stichprobengröße geschätzt wird. Also spiegelt die Unsicherheit der Parameter unser Unvermögen wider, diese Werte genau zu bestimmen; Beispiele sind Werte zur Wirksamkeit der Intervention, zum Ressourcenverbrauch und zu deren Bewertung [98]. Diese Art von Unsicherheit wird oft durch den Gebrauch von Konfidenzintervallen oder anderen statistischen Ansätzen beschrieben, welche die beobachtete Unsicherheit des

jeweiligen Parameters in der Datenquelle widerspiegeln. Da es in der Praxis selten möglich ist, die Stichproben zu vergrößern, um diese Unsicherheit um die Parameterwerte zu reduzieren, ist es wichtig, ihren Einfluss auf die Ergebnisse zu untersuchen.

Die Charakteristika der Patientenpopulation wie Alter, Geschlecht, relevante klinische Vorgeschichte sind sehr wichtig hinsichtlich der Verwendbarkeit der Ergebnisse aus der ökonomischen Evaluation. Obwohl die Patientencharakteristika in der Untersuchungspopulation variieren können, impliziert dies nicht *per se* Unsicherheit [91]. Indes kann Unsicherheit aufgrund unvollständiger Information über diese Charakteristika vorliegen. In diesem Fall sollte dies auch zusammen mit anderen Parameterunsicherheiten quantifiziert werden.

Unsicherheiten bei den Modellparametern müssen berücksichtigt werden und ihr Einfluss auf die modellierten Ergebnisse unter Verwendung geeigneter Methoden quantifiziert werden [25,91,99]. Die drei gebräuchlichsten Methoden in gesundheitsökonomischen Evaluationen sind die univariate deterministische, die multivariate deterministische und die multivariate probabilistische Sensitivitätsanalyse. Es existieren auch andere Methoden [93,100], welche aber selten in Health Technology Assessments angewandt werden.

Die univariate Sensitivitätsanalyse bezieht sich auf die Spannweite der Outcomes, die durch Variierung eines der Modellparameter innerhalb eines plausiblen Unsicherheitsbereichs resultieren, während alle anderen Parameter konstant in ihren ursprünglich ermittelten Werten gehalten werden (hierbei ist zu beachten, dass Korrelationen zwischen Inputvariablen nicht berücksichtigt werden könnten). Dies kann einen Anhaltspunkt zum Ausmaß des Einflusses dieser Unsicherheit geben, aber offensichtlich nicht die gesamte Unsicherheit der Ergebnisse quantifizieren. Univariate Sensitivitätsanalysen sind extrem hilfreich, wenn der Entscheidungsträger Annahmen über einen wichtigen Parameter trifft, die sich von der Ausgangsanalyse unterscheiden.

Mittels multivariater Sensitivitätsanalysen wird versucht, die Bewertung der Unsicherheit vollständiger zu gestalten, indem mehrere Parameter gleichzeitig verändert werden und ihr gemeinsamer Einfluss untersucht wird. Obwohl dies bei der Untersuchung von Szenarien sehr hilfreich sein kann, wenn die Unsicherheit im Vordergrund steht, ist dieses Vorgehen

zeitraubend, wenn jeder Parameter eine breite Spannweite aufweist, und nicht sehr genau, wenn Korrelationen unter den Parameterwerten unberücksichtigt bleiben.

In der probabilistischen Sensitivitätsanalyse wird jedem Parameter, der mit nicht-interessierender Unsicherheit behaftet ist, eine Verteilung zugeordnet und aus diesen Verteilungen wird simultan eine Stichprobe gezogen, um ein Set von Eingangswerten für einen Modelldurchlauf zu erhalten [34]. Dieser Vorgang wird viele Male wiederholt und der resultierende Ergebnissatz ergibt einen Eindruck des gesamten Einflusses der Parameterunsicherheit. Für eine sinnvolle Wiedergabe der Unsicherheit müssen die ausgewählten Verteilungen präzise die Unsicherheit um die Parameterschätzer darstellen und die Korrelation untereinander muss korrekt erfasst sein. Anderenfalls können Parameterkombinationen, die in Wirklichkeit selten oder nie auftreten, überrepräsentiert werden.

11.3 Empfehlung

Variabilität, die von Interesse für den Forscher ist, sollte natürlich untersucht werden. Parameterunsicherheit sowie andere Unsicherheiten, die nicht im Vordergrund stehen, aber nicht reduzierbar sind, sollten quantifiziert werden. Alle diesbezüglich unternommenen Analysen sollten hinsichtlich der Spannweiten für die eingesetzten Parameterwerte und getroffenen Annahmen vollständig dokumentiert werden. Es wird empfohlen, univariate Sensitivitätsanalysen nicht durch multivariate oder probabilistische Sensitivitätsanalysen zu ersetzen, sondern diese zu ergänzen. Für univariate und multivariate Analysen sollten die Ergebnisse sowohl in tabellarischer Form als auch mittels eines Tornado-Diagramms dargestellt werden, indem die entsprechenden Inputparameterspannweiten den Ergebnisspannweiten gegenübergestellt werden. Im Falle probabilistischer Sensitivitätsanalysen sollten die Ergebnisse als kumulative Kostenverteilungen präsentiert werden. Kosten-Effektivitäts-Akzeptanzkurven [101] werden in den IQWiG-Evaluationen nicht verlangt, da bei IQWiG-Evaluationen keine Schwellenwerte von Kosten-Effektivitäts-Verhältnissen zum Einsatz kommen. Strukturelle Sensitivitätsanalysen sollten zur Erueierung des Einflusses einer Variierung der strukturellen Annahmen durchgeführt werden.

12 Validierung

Die unterschiedlichen Ansätze zur Validierung sind auf viele Weisen klassifiziert worden. Anstatt sich auf ein bestimmtes Validierungsvorgehen und seine Terminologie zu fokussieren, soll hier geklärt werden, aus welchen Anteilen die Validierung besteht und was von jedem dieser Teile abgedeckt wird. Ein Schlüsselement befasst sich mit der Frage, ob das Modell Sinn ergibt. Die Plausibilität (Augenscheinvalidität) bezieht sich auf das Einflussdiagramm, das Modellkonzept, die Datenakquise, die funktionalen Beziehungen und die Auswahl der Modellierungstechnik. Sie müssen alle mit dem vorhandenen Wissen, der besten verfügbaren Information und der „best practice“ für jede Komponente vereinbar sein. Ein zweites Schlüsselement der Validierung setzt sich mit der korrekten Implementation des Modells auseinander (wird auch als *technische* Validität bzw. „Verifikation“ bezeichnet). Dieser Aspekt bezieht sich auf die Frage, ob das Modell in der intendierten Weise funktioniert und die Programmlogik mit einer möglichst geringen Anzahl an Programmierfehlern korrekt implementiert ist. Der dritte ausschlaggebende Aspekt ist die externe Validierung des Modells, d. h. der Nachweis, dass das Modell im Rahmen seiner definierten Limitationen korrekt die „Wirklichkeit“ wiedergibt. Dies ist von besonderer Wichtigkeit, weil ein technisch perfektes, von Programmierfehlern freies Modell mit starker Plausibilität immer noch von geringem Wert für die Entscheidungsfindung sein kann, wenn es nicht richtig das System und sein Verhalten wiedergibt.

Ein Simulationsmodell, das für einen Zweck valide ist, kann für einen anderen nicht valide sein [102]. Der externe Validierungsprozess muss daher jede intendierte Verwendung des Modells abdecken und bei Einsatz desselbigen für andere Zwecke muss die Validierung erneut erfolgen.

Es existiert kein vorgeschriebenes Vorgehen zur Validierung eines Modells, aber es gibt einige grundlegende Schritte, die eingehalten werden müssen [103].

Das Design und die Struktur des Modells sollten von anderen Sachverständigen auf dem das Modell betreffenden Gebiet überprüft werden. Dies sollte für das Einflussdiagramm und das Modellkonzept geschehen. Das Einflussdiagramm stellt einen wichtigen Aspekt der Plausibilitätsprüfung (Augenscheinvalidität) dar. Dadurch, dass es das Verständnis des Modellierers über das Krankheitsbild und seine Behandlung aufzeigt, liefert es ein

anschauliches Instrument für klinische und andere Experten zur Bewertung des geplanten Modells.

Alle in das Modell eingehenden Gleichungen sollten separat getestet werden, bevor sie in das Modell eingefügt werden. Dies wird zur Sicherstellung ihrer korrekten Wiedergabe und der richtigen Ergebnisgenerierung für die eingegebenen Werte unternommen. Die Gleichungen sollten auch anhand von Extremwerten, die sich an den Modellrändern orientieren, durchgerechnet werden, um sicherzustellen, dass keine Probleme in diesen Bereichen auftreten.

Das Testen aller Modellaspekte nach seiner Vollendung ist aufgrund der Modellkomplexitäten und auftretenden Interaktionen ein umfangreiches Unterfangen. Der einzige sinnvolle Ansatz ist die permanente Überprüfung des Modells während seines Entwicklungsprozesses. Dies kann am besten gewährleistet werden, indem das Modell modular aufgebaut wird. Jedes Modul wird nach seiner Entwicklung getestet. Die modulare Überprüfung bezieht sich auf einen korrekten Output bei gegebenen Inputs für das Modul (z. B. Entitäten und Dateneingabe im Falle einer Diskreten Ereignissimulation). Diese Überprüfung muss auf die Einbeziehung von Extremwerten für die Inputs und extremen Konditionen erweitert werden, um sicherzustellen, dass das Modell auch unter solchen Bedingungen korrekt funktioniert. Jede Änderung in einem bereits überprüften Modul erfordert eine erneute Überprüfung dieses Moduls. Während neue Module hinzugefügt werden, sollten sie kontinuierlich überprüft werden, um ein konsistentes Verhalten zu gewährleisten. In einer Simulation auf individuellem Level (z. B. Markov-Modell mit Monte-Carlo-Simulation erster Ordnung oder Diskrete Ereignissimulation), besteht die grundlegende Methode darin, einzelne, klar definierte Patienten zu testen, um sicherzugehen, dass die Patienten die erwarteten Pfade durchlaufen und die erwarteten Ereignisse erleben. Darüber hinaus sollte dieselbe Vorgehensweise, mit welcher die Einzelmodule getestet wurden, für das gesamte Modell angewandt werden.

Ein guter Ansatz zur externen Validierung des Modells ist, das Modell unter Verwendung von Inputwerten laufen zu lassen, die aus einer oder mehreren historischen Situationen wie beispielsweise bereits durchgeführten klinischen Studien, Registern oder Outcomestudien stammen, aus welchen bislang keine Daten in das getestete Modell eingeflossen sind. Die Modellergebnisse werden dann mit den bereits bekannten Ergebnissen verglichen [9]. Durch

Überprüfung von Studien, die nicht als Quellen für Inputwerte verwendet wurden, wird die Fähigkeit des Modells, die Realität wiederzugeben, geprüft. Diese Überprüfung muss allerdings relativiert werden, da Studien selbst substanzielle Abweichungen von der Realität darstellen. Idealerweise müsste der gleiche Prozess mit Daten aus der gegenwärtigen Versorgung durchlaufen werden, jedoch werden diese nur in seltenen Fällen verfügbar sein.

13 Modelldokumentation

Ein detaillierter technischer Bericht mit Beschreibung aller Modellierungsschritte von der Entwicklung des Einflussdiagramms bis zur finalen Validierung ist zu erstellen. Zusätzlich ist eine voll ausführbare Version des Modells einschließlich eines Anwendermanuals zur Verfügung zu stellen. In Anlehnung an andere vorgeschlagene Leitlinien [5,12,20,25,99] sollte die Dokumentation des Modells Folgendes beinhalten:

- Das zur Anleitung der Modellentwicklung verwendete Einflussdiagramm
- Details zum Modellkonzept
 - Beschreibung der Zielpopulation(en) im Rahmen der Evaluation einschließlich Subgruppen
 - Beschreibung der evaluierten Gesundheitstechnologien
 - Auswahl und Begründung der Modellsettings (Simulationsgröße, Zeithorizont, Diskontierungsraten etc.)
 - Review zu gesundheitsökonomischen Evaluationen im untersuchten Therapiebereich
- Beschreibung aller Datenquellen. Eine Begründung der Auswahl von Datenquellen muss gegeben werden.
- Details zu allen im Modell eingesetzten funktionalen Beziehungen. Wenn diese speziell für das Modell entwickelt wurden, müssen detaillierte Angaben zu den angewandten Methoden gegeben werden.
- Auflistung aller Annahmen hinsichtlich der Datenquellen und der Modellstruktur. Besonders wichtig ist eine detaillierte Darstellung jeglicher Annahme und verwendeten Technik zur Projektion über den Zeitraum hinaus, für welchen die Daten gelten.
- Begründung für die verwendete Modellierungstechnik

- Beschreibung, in welcher Art die Technik den erforderlichen Eigenschaften (siehe Kapitel 7) entspricht
- Überblick zu den eingesetzten Validierungstechniken und ihrer Ergebnisse
- Detaillierte Ergebnisdarstellung einschließlich einer Bewertung des Einflusses von:
 - Anwendung der Gesundheitstechnologie in relevanten Populationssubgruppen
 - Unsicherheit der Inputdaten
 - Sollten probabilistische Sensitivitätsanalysen durchgeführt worden sein, ist eine Spezifizierung folgender Punkte einzuschließen:
 - die eingesetzten Wahrscheinlichkeitsverteilungen und ihre Quellen
 - Korrelationen zwischen Inputparametern
 - jegliche strukturelle Varianten
 - Interpretation der Ergebnisse einschließlich einer Beschreibung der Limitationen des angewandten Ansatzes

Eine elektronische Version des Modells muss zur Verfügung gestellt werden mit dem Einverständnis, dass das Modell öffentlich zugänglich gemacht wird und ggf. für zukünftige Evaluationen adaptiert werden kann. Die elektronische Modellversion muss vollkommen zugänglich sein und den Reviewern sowie der Öffentlichkeit ermöglichen, alle in der Analyse eingesetzten Formeln und Beziehungen einzusehen und das Modell mit unterschiedlichen Inputdaten laufen zu lassen. Um das Review des Modells zu erleichtern, sollte der elektronischen Version eine Anwenderanleitung beigelegt werden, die beschreibt, welche Software und Hardware benötigt werden, wie Modellinputs verändert werden können, wo

diese Inputs im Modell gefunden werden können, wie das Modell ausgeführt werden kann und wie Ergebnisse extrahiert werden können.

Literaturverzeichnis

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. 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.
4. 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.
5. 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.
6. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; 15: 1295-1310.
7. 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.
8. 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.
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. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Cost-effectiveness analysis. In: *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2000. S. 98-102.
16. Nuijten MJC. The selection of data sources for use in modelling studies. *Pharmacoeconomics* 1998; 13(3): 305-316.
17. 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.
18. 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. S. 247-255.
19. 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.
20. Soto J. Health economic evaluations using decision analytic modeling. *Int J Technol Assess Health Care* 2002; 18(1): 94-111.
21. Sculpher MJ, Drummond MF. Analysis sans frontières: can we ever make economic evaluations generalisable across jurisdictions? *Pharmacoeconomics* 2006; 24(11): 1087-1099.
22. Siebert U. Entscheidungsanalytische Modelle zur Sicherung der Übertragbarkeit internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems: ein Methodenbeitrag zu HTA. Köln: DAHTA@DIMDI; 2005. (HTA Bericht; Band 16). URL: http://gripsdb.dimdi.de/de/hta/hta_berichte/hta099_bericht_de.pdf.
23. 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.
24. 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.
25. National Institute for Clinical Excellence. Guide to the methods of technology appraisal. London: NICE; 2004. URL: http://www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf.
26. 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.
27. Agency for Healthcare Research and Quality. Registries for evaluating patient outcomes: a user's guide; AHRQ publication no. 07-EHC001-1 [Online]. 04.2007 [Zugriff am 22.06.2008]. URL: <http://effectivehealthcare.ahrq.gov/repFiles/PatOutcomes.pdf>.
28. Statistisches Bundesamt Deutschland [Online]. 2007 [Zugriff am 01.07.2007]. URL: <http://www.destatis.de/jetspeed/portal/cms>.

-
29. The Human Mortality Database [Online]. 2007 [Zugriff am 01.06.2007]. URL: <http://www.mortality.org>.
30. 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.
31. Drummond MF, Sculpher MJ. Common methodological flaws in economic evaluations. *Med Care* 2005; 43(7 Suppl): 5-14.
32. 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: DHA; 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).
33. 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. S. 234-235.
34. Gold MR, Russell LB, Siegel JE (Ed). *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. Nixon RM, Thompson SG. Parametric modelling of cost data in medical studies. *Stat Med* 2004; 23(8): 1311-1331.
41. 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.
42. 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.

-
43. 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.
44. 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. S. 95-134. (The IMA volumes in mathematics and its applications; Band 116).
45. 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. S. 2191-2230.
46. 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.
47. Mullahy J. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *J Health Econ* 1998; 17(3): 247-281.
48. Kleinbaum DG, Klein M. *Survival analysis: a self-learning text*. New York: Springer; 1996.
49. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Methoden: Version 3.0* [Online]. 27.05.2008 [Zugriff am 25.08.2008]. URL: http://www.iqwig.de/download/IQWiG_Methoden_Version_3_0.pdf.
50. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer; 2001.
51. Draper NR, Smith H. *Applied regression analysis*. New York: Wiley; 1998.
52. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999; 10(1): 37-48.
53. Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci USA* 2002; 99(Suppl 3): 7280-7287.
54. Cooper K, Brailsford SC, Davies R. Choice of modeling technique for evaluating health care interventions. *J Oper Res Soc* 2007; 58(2): 168-176.
55. 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.
56. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. *Pharmacoeconomics* 2008; 26(2): 131-148.
57. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. *Pharmacoeconomics* 2006; 24(11): 1043-1053.

-
58. Caro JJ. Pharmacoeconomic analyses using discrete event simulation. *Pharmacoeconomics* 2005; 23(4): 323-332.
59. 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.
60. Koopman J. Modeling infection transmission. *Annu Rev Public Health* 2004; 25: 303-326.
61. 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. S. 141-171.
62. Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required? *Value Health* 2006; 9(4): 244-252.
63. Briggs AH, Claxton K, Sculpher M. *Decision modelling for health economic evaluation: handbooks for health economic evaluation*. Oxford: Oxford University Press; 2006.
64. 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.
65. Lee RC, Donaldson C, Cook LS. The need for evolution in healthcare decision modeling. *Med Care* 2003; 41(9): 1024-1033.
66. Banks J, Carsons J.S., Nelson BL, Nicol DM. *Discrete-event system simulation*. New Jersey: Pearson Prentice Hall; 2005.
67. Macal CM, North MJ. Tutorial on agent-based modeling and simulation [Online]. 2005 [Zugriff am 23.06.2008]. URL: http://www.dis.anl.gov/publications/articles/Macal_and_North_ABMS_Tutorial.pdf.
68. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making* 1983; 3(4): 419-458.
69. Briggs AH, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998; 13(4): 397-409.
70. Siebert U, Mühlberger N, Schöffski O. Evidenzsynthese: Meta-Analysen und Entscheidungsanalysen. In: Schöffski O, Graf von der Schulenburg JM (Ed). *Gesundheitsökonomische Evaluationen*. Berlin: Springer; 2007. S. 261-310.
71. 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.
72. Fishman GS. *Discrete-event simulation: modeling, programming, and analysis*. Berlin: Springer; 2001.

-
73. Bratley P, Fox BL, Schrage LE. A guide to simulation. New York: Springer; 1987.
74. Kelton WD, Sadowski RP, Sturrock DT. Simulation with ARENA. Boston: McGraw-Hill; 2007.
75. 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. In: Peters BA, Smith JS, MEdeiros DJ, Rohrer MW. Proceedings of the 2001 Winter Simulation Conference. S. 1386-1391. 2001.
76. Cahill W, Render M. Dynamic simulation modeling of ICE bed availability. In: Farrington PA, Nembhard HB, Sturrock DT, Evans GW (Ed). Proceedings of the 1999 Winter Simulation Conference. S. 1573-1576.
77. 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.
78. 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.
79. Proctor T. Simulation in hospitals. Health Manpow Manage 1996; 22(5): 40-44.
80. Davies HTO, Davies R. Simulating health systems: modelling problems and software solutions. Eur J Oper Res 1995; 87(1): 35-44.
81. 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.
82. Siebert U. Transparente Entscheidungen in Public Health mittels systematischer Entscheidungsanalyse. In: Schwartz FW, Badura B, Busse R, Leidl R, Raspe H, Siegrist J et al. (Ed). Das Public Health Buch: Gesundheit und Gesundheitswesen. München: Urban & Fischer; 2003. S. 485-502.
83. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making 1993; 13(4): 322-338.
84. Duffie D, Glynn P. Estimation of continuous-time Markov processes sampled at random time intervals. Econometrica 2004; 72(6): 1773-1808.
85. Stewart WJ. Introduction to the numerical solution of Markov chains. Princeton: Princeton University Press; 1994.
86. 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 MIE. Amsterdam: IOS Press; 2006. S. 699-704.

-
87. 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.
88. 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.
89. Anderson RM, May RM. *Infectious disease of humans: dynamics and control*. Oxford: Oxford University Press; 1991.
90. Bojke L, Claxton K, Palmer S, Sculpher MJ. Defining and characterising structural uncertainty in decision analytic models. *CHE Research Paper 9* [Online]. 2006 [Zugriff am 22.06.2008]. URL: <http://www.york.ac.uk/inst/che/pdf/rp9.pdf>.
91. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000; 17(5): 479-500.
92. 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.
93. Law AM, Kelton WD. *Simulation modelling and analysis*. New York: McGraw Hill; 2000.
94. Ross SM. *Simulation*. Oxford: Elsevier Academic Press; 2006.
95. Cantor SB, Ganiats TG. Incremental cost-effectiveness analysis: the optimal strategy depends on the strategy set. *J Clin Epidemiol* 1999; 52(6): 517-522.
96. Drummond MF, McGuire AE. *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001.
97. Karlsson G, Johannesson M. Cost-effectiveness analysis and capital costs. *Soc Sci Med* 1998; 46(9): 1183-1191.
98. 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.
99. Collège des Économistes de la Santé. French guidelines for the economic evaluation of health care technologies [Online]. 09.2004 [Zugriff am 26.08.2008]. URL: http://www.ces-asso.org/docs/France_Guidelines_HE_Evaluation.PDF.
100. Oberkampf WL, DeLand SM, Rutherford BM, Diegetd KV, Alvine KF. Error and uncertainty in modeling and simulation. *Reliability Engineering & System Safety* 2002; 75(3): 333-357.
101. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves-facts, fallacies and frequently asked questions. *Health Econ* 2004; 13(5): 405-415.

102. Law AM, McComas MG. How to build valid and credible simulation models [Online]. 2001 [Zugriff am 23.06.2008]. URL: <http://www.informs-sim.org/wsc01papers/004.PDF>.

103. Sargent RG. Validation and verification of simulation models [Online]. 2004 [Zugriff am 23.06.2008]. URL: <http://portal.acm.org/citation.cfm?id=1161734.1161742>.