

Arbeitspapier Modellierung

Version 1.0 vom 12.10.2009

Kontakt:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)
Dillenburger Straße 27
D-51105 Köln

Tel.: +49-221/35685-0

Fax: +49-221/35685-1

E-Mail: knb-methoden@iqwig.de

Inhaltsverzeichnis

Abkürzungsverzeichnis	iii
1 Hintergrund	1
2 Definition des Begriffs „Modell“	2
3 Ziel einer gesundheitsökonomischen Modellierung	3
4 Vorgehen bei der Modellentwicklung	4
5 Einflussdiagramme	6
5.1 Definition und Zweck.....	6
5.2 Komponenten	6
6 Modellkonzept	7
6.1 Definition und Zweck.....	7
6.2 Komponenten	7
6.3 Validierung.....	8
7 Modelleigenschaften	9
7.1 Hinreichende Detailliertheit	9
7.2 Perspektive	9
7.3 Zeithorizont	10
7.4 Diskontierung.....	10
8 Datenbewertung	11
8.1 Ziele.....	11
8.2 Potenzielle Datenquellen.....	12
8.2.1 Klinische Studien	12
8.2.2 Epidemiologische Studien.....	13
8.2.3 Prozessdatenbanken	13
8.2.4 Register.....	14
8.2.5 Kompilierte Statistiken.....	14
8.2.6 Expertenmeinung	15
9 Funktionale Beziehungen	16
9.1 Definition	16
9.2 Zweck.....	16
9.3 Allgemeine Konzepte	16
10 Modellierungstechniken	19
10.1 Die Wahl der Modellierungstechnik	19

10.2	Aspekte bei der Wahl der Modellierungstechnik.....	20
10.2.1	Kohortenmodelle versus Individualmodelle	20
10.2.2	Ereignisbasierte versus zustandsübergangsbasierte Simulation.....	21
10.3	Techniken.....	22
10.3.1	Entscheidungsbaumanalysen.....	22
10.3.2	Diskrete Ereignissimulation	22
10.3.3	Markov-Modelle	25
10.3.4	Agentenbasierte Simulation	28
10.3.5	Transmissionsmodelle.....	29
11	Umgang mit Variabilität und Unsicherheit	31
11.1	Arten von Variabilität.....	31
11.2	Unsicherheit und deren Quantifizierung	31
11.3	Umgang mit Unsicherheit und ihre Auswirkungen auf die Ergebnisse.....	32
12	Validierung	34
13	Modelldokumentation.....	36
	Literaturverzeichnis.....	38

Abkürzungsverzeichnis

G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
MTC	Mixed Treatment Comparison
NYHA	New York Heart Association
RCT	Randomized Controlled Trial
ROC	Receiver Operating Characteristic

1 Hintergrund

Hochwertige Daten zu den Gesundheitseffekten einer Therapie werden in klinischen Studien, idealerweise in randomisierten kontrollierten Studien (RCTs), ermittelt. Diese fließen auf der Nutzenseite in eine Kosten-Nutzen-Bewertung ein. Ökonomische Daten werden in klinischen Studien meist nicht erhoben. Sofern dies doch der Fall ist, sind diese Daten alleine häufig nicht ausreichend für eine vollständige und fundierte Aufbereitung der Kosten einer Gesundheitstechnologie. Klinische Studien liefern selten Informationen zu den langfristigen ökonomischen Konsequenzen, die mit der Einführung einer neuen Technologie einhergehen. Auch greifen sie nicht immer adäquat und vollständig die für die Kostenseite in Deutschland relevanten Versorgungsaspekte ab [1,2]. Weiterhin kann auch ein protokollinduzierter Ressourcenverbrauch im Rahmen klinischer Studien zu Fehleinschätzungen auf der Kostenseite führen. Aus diesen Gründen ist die Modellierung ökonomischer Effekte einer Gesundheitstechnologie eine wesentliche Komponente der gesundheitsökonomischen Evaluation.

Wenn sich an die Nutzenbewertung, die vor der ökonomischen Bewertung erfolgt, eine Modellierung anschließt, können Modellierungstechniken zur Schätzung sowohl der Kosten als auch der gesundheitlichen Konsequenzen eingesetzt werden. Dies kann 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 zusammenfassende Maße), oder der medizinische Nutzen in eine Kardinalskala transformiert werden muss. Darüber hinaus können durch Modellierung Vorhersagen darüber getroffen werden, wie sich die aus klinischen Studien ermittelten Nutzeneffekte unter anderen Bedingungen oder Zeitperspektiven darstellen würden. Hierbei können prognostische Anpassungen vorgenommen werden. Da die Vorhersagen mit großer Unsicherheit behaftet sind, ist darauf zu achten, dass kein neuer Nutzen generiert wird, z. B. durch die Einbeziehung neuer Endpunkte oder die Umkehrung der aus Studien gewonnenen Ergebnisse.

2 Definition des Begriffs „Modell“

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 (zuschreibbare Modellergebnisse) bei einem gegebenen Satz von Inputparametern zu schätzen und die Auswirkungen von Veränderungen der Inputparameter auf das System zu modellieren [3]. Daraus ergibt sich, dass jede über die direkte Anwendung beobachteter Daten hinausgehende Evaluation als Modell betrachtet werden kann [4]. Selbst die statistische Auswertung von Daten beinhaltet bis zu einem gewissen Ausmaß Aspekte einer Modellierung, beispielsweise im Fall von Meta-Analysen klinischer Studien. Selbstverständlich können Modelle die Wirklichkeit nicht perfekt abbilden: Sie basieren auf einem reduzierten Satz von Komponenten und erfordern vereinfachende Annahmen [5]. 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 gesundheitsökonomische 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 and whose purpose is to estimate the effects of an intervention on valued health consequences and costs“ [6]. Eine aktuelle Veröffentlichung zum Thema „Modellierungstechniken“ gibt eine ähnliche Definition: „a formal quantified comparison of health technologies synthesizing sources of evidence on costs and benefits“ [7]. Beide Definitionen sind vereinbar mit den Anforderungen an gesundheitsökonomische Evaluationen, die vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) durchgeführt werden.

3 Ziel einer gesundheitsökonomischen Modellierung

Ziel einer gesundheitsökonomischen Modellierung ist die Ermittlung von Erwartungswerten für die klinischen und ökonomischen Effekte der verglichenen Interventionen. Die Modellierung im Rahmen einer gesundheitsökonomischen Bewertung des IQWiG stellt auf diese Weise Informationen bereit, die zur Einschätzung der Effizienz einer Intervention für die Festlegung von Höchstbeträgen benötigt werden [8]. Das vorliegende Arbeitspapier gibt Handlungsempfehlungen für das Vorgehen bei der Modellentwicklung, die Wahl der Modellierungstechnik, die Berücksichtigung von Unsicherheit und Variabilität, die Validierung des Modells und die Form, in der Modellierungsstudien und die diesen zugrunde liegenden Annahmen dokumentiert und beschrieben werden müssen.

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

4 Vorgehen bei der Modellentwicklung

Eine Beschreibung des Vorgehens bei der Modellentwicklung und Wahl der Modellierungstechnik sowie bei der Validierung, Analyse und Berichterstattung findet sich im anschließenden Teil dieses Arbeitspapiers. 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/en ist / sind relevant für die Analyse?
 - c. Welches sind die relevanten Kostenkategorien, die mit der betreffenden Indikation und den evaluierten Interventionen in Zusammenhang stehen?
2. Entwurf eines Einflussdiagramms, welches grafisch die verschiedenen Faktoren darstellt, die einen Einfluss auf die zu modellierende/n Fragestellung/en haben können, und mögliche Zusammenhänge zwischen diesen aufzeigt [6].
3. Eng in Zusammenhang mit der Entwicklung des Einflussdiagramms steht die Präzisierung des Modellkonzepts, welches als Konstruktionsplan für das Modell dient. Das Modellkonzept soll darstellen, wie die Gesundheitseffekte der evaluierten Handlungsalternativen mit den konkreten Gesundheitseffekten und dem Ressourcenverbrauch zusammenhängen, sowie deren Projektion über den Zeithorizont des Modells verdeutlichen. Außerdem soll das Konzept aufzeigen, 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 Forschungsfrage/n verwendeten Modellierungstechnik sein (und nicht umgekehrt – d. h. es sollte nicht zuerst die Modellierungstechnik festgelegt werden).
4. Nachdem das Modellkonzept erstellt worden ist, sollten die Daten gesammelt werden, mit denen das Modell versehen werden kann. Es entspricht nicht der guten Praxis, das Modell um die verfügbaren Daten herum zu entwerfen, sondern das Modell sollte aus dem Einflussdiagramm und dem Modellkonzept abgeleitet werden. Sofern in einzelnen Bereichen mit der systematischen Datenrecherche keine Daten zu Kosten gefunden werden können, wird eine gezielte Suche in zusätzlichen Datenquellen oder die Primärerhebung von Daten erwogen. Dies kann beispielsweise in Form von spezifischen Beobachtungsstudien erfolgen oder durch die zielgerichtete Auswertung bestehender Register oder Datenbanken, welche Informationen beinhalten, die ursprünglich für andere Zwecke gesammelt wurden (siehe z. B. Abschnitt 8.2.4).

5. Auf die 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 aus der Verfügbarkeit der Daten und der gewählten Modellierungstechnik hergeleitet. Hierbei ist es entscheidend, dass die 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.
6. 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 sein [9].
7. Erst nach der Validierung des Modellkonzepts, der Auswahl der Daten, der Entwicklung der verwendeten funktionalen Zusammenhänge und der Wahl der Modellierungstechnik sollte man mit der Implementierung und Programmierung des Modells fortfahren.
8. Mit der technischen Validierung sollte bereits während der Implementierung begonnen werden. Sie wird aber erst nach Abschluss der Programmierung zusammen mit der Plausibilitätskontrolle abgeschlossen.
9. 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].
10. Der Bericht zur Modellierung sollte alle oben beschriebenen Prozessschritte beinhalten und deren Ergebnisse zusammenfassen. Berichte müssen von einem vollständig funktionsfähigen und zugänglichen elektronischen Modell begleitet werden. Das Modell muss öffentlich zur Verfügung gestellt werden, um Transparenz zu gewährleisten und das Vertrauen in dessen Integrität zu stärken.

Die Entwicklung eines validen Modells ist ein beträchtliches Unterfangen, weshalb zur Verwendung von bereits validierten Modellen (siehe Kapitel 12) 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 notwendig sein.

5 Einflussdiagramme

5.1 Definition und Zweck

Einflussdiagramme werden zunehmend bei der Entwicklung von ökonomischen Modellen eingesetzt [11]. Ein Einflussdiagramm bildet grafisch die wesentlichen Zusammenhänge des Modells ab. Es zeigt die Faktoren auf, die Einfluss auf eine zu modellierende Fragestellung bzw. Fragestellungen haben oder haben können. Hierdurch ermöglicht das Diagramm allen Interessenten, auch ohne Verständnis der technischen Aspekte des Modells an sich, die Idee des Modellierers zu verstehen und Änderungsempfehlungen abzugeben.

Trotz seines Namens ist das Einflussdiagramm nicht per se eine kausale Darstellung. Vielmehr legt es die Beziehungen zwischen den Faktoren offen, ohne notwendigerweise stochastische Abhängigkeitsverhältnisse zwischen diesen zu berücksichtigen.

5.2 Komponenten

Das Einflussdiagramm muss alle wichtigen Aspekte der modellierten Erkrankung sowie derer wechselseitigen Beziehungen enthalten, einschließlich:

- Patientencharakteristika
- Pathophysiologie und klinischer Verlauf der Erkrankung
- Behandlung und Management der Erkrankung
- Effekte einer Intervention

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).

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 und Verbindung erklärt sein. Darüber hinaus sollte eine Begründung für das Weglassen möglicher Verbindungen abgegeben werden.

6 Modellkonzept

6.1 Definition und Zweck

Das Modellkonzept bildet präzise und detailliert die analytische Struktur und den Ansatz des Modells ab. Es ergibt sich aus dem Einflussdiagramm und stellt das beabsichtigte Design in deutlich größerer Tiefe dar.

Die Erarbeitung des Modellkonzepts ist ein kritischer Schritt in der Modellentwicklung und sollte daher gut dokumentiert und beschrieben werden. Das Modellkonzept soll die Nachvollziehbarkeit und die Validität des späteren Modells erhöhen. Es soll verdeutlichen, wie das modellierte Indikationsgebiet und die Effekte der Interventionen verstanden und bewertet werden. Alle Annahmen zu den Beziehungen zwischen den Modellkomponenten sollten explizit genannt werden. Da auch die differenziertesten Modelle Vereinfachungen der Realität mit erforderlichen Annahmen und Einschränkungen in Bezug auf eingeschlossene Inhalte [10,12] darstellen, kann das Modell nur dann richtig verstanden werden, wenn das Modellkonzept präzisiert und gut dokumentiert wird.

6.2 Komponenten

Die Präzisierung des Modellkonzepts sollte folgende Aspekte berücksichtigen:

- Welche klinischen Ereignisse sollten einbezogen werden und wodurch wird deren Ausprägung beeinflusst?
 - Können diese Ereignisse wiederholt auftreten? Wodurch wird das Risiko einer Wiederholung beeinflusst?
 - Beeinflusst das Auftreten 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 vom IQWiG ermittelte patientenrelevante therapeutische Zusatznutzen der Intervention/en 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 Patienten / medizinischen Versorger die Gesundheitseffekte einer Maßnahme?
 - Welche Bedeutung hat die Adhärenz der Patienten / Leistungserbringer und wie werden zukünftige medizinische Entscheidungen hierdurch beeinflusst?
 - Ist die Interaktion zwischen Patienten relevant, wie beispielsweise im Fall von akuten Infektionserkrankungen?
- Wie werden die Kosten durch das Patientenmanagement, die Anamnese, das Patientenverhalten und den Zeitpunkt bzw. den Schweregrad von Komplikationen beeinflusst?

6.3 Validierung

Modelle sollten „consistent both with a coherent theory of the health condition being modeled and with available evidence regarding causal linkages between variables“ [6] sein. Deshalb kann die Fertigstellung des zentralen Modellkonzepts nicht ohne eine Validierung erfolgen.

Zur Gewährleistung der Validierung des Modellkonzepts ist es wichtig, eine vollständige Dokumentation der zur Definition des Modellkonzepts verwendeten Informationen und der hypothetischen Verbindungen einzuschließen.

Auch ist es hilfreich, das Konzept mit anderen Modellkonzepten 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 Indikationsgebiet zur Anwendung kommen, beinhalten.

Übereinstimmung mit publizierten Modellen allein ist jedoch angesichts deren unterschiedlicher Qualität und des Fehlens eines „Gold Standard“ für die Komponenten eines „guten Modells“ [13] unzureichend. Wesentlich ist, dass Modelle für Personen (z. B. anerkannte Experten, Kliniker, Patienten, Angehörige), die sich mit einer Erkrankung auskennen, Sinn ergeben [14]. Zur Validierung des Konzeptes müssen daher u. a. anerkannte Experten einbezogen werden, die sich mit den modellierten Indikationsgebieten auskennen.

7 Modelleigenschaften

Modelle zur Bewertung von Nutzen und Kosten einer Technologie im Vergleich zu Therapiealternativen müssen 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 von den Modelleigenschaften abhängt, sondern diese auch durch Peer-Review-Verfahren und vollständig zugängliche Modelle überprüft bzw. überprüfbar gemacht wird.

7.1 Hinreichende Detailliertheit

Die Modelle des IQWiG müssen detailliert die Kosten und den Nutzen abbilden, die sich für die zu betrachtende Gesundheitstechnologie ergeben. Dies erfordert folgende Informationen:

- eindeutige Bestimmung und Eingrenzung des Zusatznutzens, der mit der Intervention erzielt wird,
- Erfassung möglichst aller Krankheits- und Behandlungsaspekte, die Einfluss auf die Kosten- oder Nutzenkomponenten des Modells haben können, z. B. in folgenden Bereichen:
 - Demografie
 - Epidemiologie
 - spezifische Behandlungsmuster
 - Behandlungssadhärenz

Wo Daten zu diesen Aspekten fehlen, ist es wichtig, den möglichen Einfluss auf die Ergebnisse durch angemessene Wertebereiche der fehlenden Daten zu explorieren.

7.2 Perspektive

Die Perspektive der Untersuchung sollte primär die Perspektive der Versichertengemeinschaft der Gesetzlichen Krankenversicherung (GKV) sein. Diese ist nicht identisch mit der Perspektive der GKV selbst, da sie Ressourcen der Versicherten einbezieht.

Sofern möglich, sollten Ergebnisse zu Kosten sowohl in aggregierter als auch in vollständig disaggregierter Form berichtet werden. Eine ausführliche Darstellung der Kostenkomponenten findet sich im Arbeitspapier zur Kostenbestimmung.

7.3 Zeithorizont

Es ist generell anzustreben, das Modell für die folgenden zwei Zeiträume aufzustellen und zu berechnen:

- Primäres Szenario: Die Modellierung beschränkt sich auf einen durch in die Nutzenbewertung eingeschlossenen Studien (in der Regel RCTs) belegten Zeitraum. Dadurch wird parallel eine Validierung des Modells über vorhandene Studiendaten ermöglicht.
- Sekundäres Szenario: Der Zeitraum der Modellierung geht über die o. g. Studiendauer hinaus, wenn dies für den Entscheidungsträger relevant ist.

7.4 Diskontierung

Alle in dem Modell evaluierten Kosten- und Nutzenkomponenten müssen in angemessen diskontierter Form berichtet werden. Sensitivitätsanalysen sollten eine weite Spanne an Diskontierungsraten abdecken, um die Auswirkungen des Diskontierungseffekts auf die Nettokosten und Gesundheitseffekte der Intervention zu untersuchen. Diskontierungsraten werden vom IQWiG festgelegt und, sofern erforderlich, aktualisiert. Detaillierte Informationen zur Diskontierung von Kosten und Nutzen finden sich im Methodenpapier [15] sowie im Arbeitspapier zur Kostenbestimmung [16].

8 Datenbewertung

Die verfügbaren Daten sollten systematisch zusammengetragen und ihre Qualität und Relevanz als Modellinputs sowie in Bezug auf die Risikofunktionen untersucht werden. Evidenzhierarchien, anhand derer die Validität der Evidenz verschiedener Studienarten eingestuft wird, werden auch in gesundheitsökonomischen Studien angewandt [17]. Dies ist sinnvoll bei der Nutzen- [18], jedoch nicht immer bei der Kostenermittlung. Randomisierte kontrollierte Studien erfassen kostenrelevante Versorgungsaspekte teilweise unvollständig [17]. Kontrollierte prospektive Studien sind zur Kosteneinschätzung den retrospektiven Studien überlegen; jedoch kann auch eine retrospektiv gebildete Datenbankkohorte oder eine Fallkontrollstudie ergänzende Kosteninformationen bereitstellen. Da allerdings nicht-randomisierte Studien nicht per se das für die Bewertung relevante Versorgungsgeschehen adäquat widerspiegeln, muss auch bei Verwendung solcher Studien sichergestellt werden, dass kostenrelevante Versorgungsaspekte entsprechend erfasst wurden.

Üblicherweise werden Daten aus verschiedenen Datenbanken gewonnen. Dazu gehören auch Studien zu Behandlungseffekten, Kohortenstudien zur Untersuchung von Parametern und Risikofaktoren für den natürlichen Krankheitsverlauf und die damit verbundene Lebenserwartung, sektorenübergreifende Fragebögen zur Erfassung von Lebensqualitätsdaten, Registerdaten zu Ressourcenverbrauch und Kosten, Prozessdatenbanken oder kompilierte 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 die Studienauswahl ist die Übertragbarkeit der Ergebnisse auf den deutschen Kontext. Eine Übertragbarkeit von Kostendaten aus anderen Gesundheitssystemen ist aufgrund von Systemunterschieden selten gegeben und nur unter sehr strengen Voraussetzungen möglich [3,19]

Die Begründung für die Auswahl bestimmter Datenquellen sollte in der Dokumentation der Modellierung enthalten sein, insbesondere für die Parameter, die die Modellergebnisse wesentlich beeinflussen. Details der Datenverarbeitungsmethoden zur Generierung der benötigten Inputs, Funktionen und Verteilungen der Parameter müssen ebenfalls dargestellt werden [13,20,21].

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 verlässliche Schätzer zu erhalten.

8.2 Potenzielle Datenquellen

8.2.1 Klinische Studien

Das IQWiG bewertet den Nutzen anhand von aus klinischen Studien erhobenen patientenrelevanten Endpunkten, insbesondere zu den Zielgrößen Mortalität, Morbidität und gesundheitsbezogene Lebensqualität sowie zu validen Surrogatparametern. Kriterien, denen ein valider Surrogatparameter genügen muss, sind im allgemeinen Methodenpapier des IQWiG beschrieben [18]. Die patientenrelevanten Nutzeneffekte präventiver, diagnostischer und therapeutischer Interventionen werden in den Nutzenbewertungen des IQWiG, basierend auf den Daten kontrollierter klinischer Studien, zur Verfügung gestellt.

Abhängig von der Qualität, Relevanz und Verfügbarkeit der Daten können im Modell Inputwerte verwendet werden, die auf den Ergebnissen einer Studie basieren oder aus gepoolten Datensätzen bzw. einer Meta-Analyse berichteter Studienergebnisse stammen. Sofern zum Zeitpunkt der Evaluation eines neuen Produktes im Hinblick auf die Erstattung nur relativ kleine klinische Studien mit kurzem Zeithorizont vorliegen, reichen diese häufig nicht zur Aufdeckung von signifikanten Unterschieden auf der Kostenseite aus. Dieser Sachverhalt stellt eine wesentliche Limitation bei ökonomischen Evaluationen dar, da diese ein Versorgungsgeschehen abbilden sollen, welches sich meist über längere Zeithorizonte als die in klinischen Studien abgebildeten erstreckt [22-24].

Zur Zusammenfassung der relativen oder absoluten Behandlungseffekte einzelner Studien werden in der Regel Meta-Analysen eingesetzt. Head-to-Head-Vergleiche sind nicht regelhaft für alle Interventionen verfügbar, wodurch indirekte Vergleiche erforderlich werden, um alle zu bewertenden Behandlungen einzubeziehen. Es sollten geeignete Methoden zur Ableitung der indirekten Schätzer der Behandlungseffekte verwendet werden [25-27]. Im Allgemeinen ist ein adäquater Ansatz durch eine sogenannte Mixed-Treatment-Comparison-(MTC)-Analyse gegeben, in der sowohl direkte als auch indirekte Evidenz zur Schätzung der Behandlungseffekte genutzt wird [28-34].

Die Nutzung von Daten aus klinischen Studien zur Kostenschätzung kann grundsätzlich problematisch sein, da zum einen Daten zu kostenrelevanten Versorgungsaspekten oftmals nicht vollständig erfasst werden und da andererseits protokollinduzierte Kosten entstehen können, die in der gesundheitsökonomischen Bewertung nicht berücksichtigt werden sollten [1,2]. Bei Daten aus multizentrischen oder internationalen Studien sollte geprüft werden, ob die hierin befolgten Versorgungspfade den für Deutschland relevanten Ressourcenverbrauch adäquat widerspiegeln. Landesspezifische Unterschiede beispielsweise in Bezug auf die Vergleichsstrategien oder die Adhärenz der Patienten / Leistungserbringer könnten einen verzerrenden Einfluss auf Kostenschätzungen haben [22-24]. Ebenfalls können epidemiologische Parameter, die die Kosten beeinflussen, beträchtlich zwischen Ländern variieren [19]. Die in diesen Studien erhobenen Daten können deshalb nicht ohne weitere Auswertung direkt auf Deutschland übertragen werden [3,35-37].

Aus den zuvor genannten Gründen oder auch bei vollständigem Fehlen von Kosteninformationen aus klinischen Studien kann die Nutzung anderer Datenquellen, beispielsweise aus Beobachtungsstudien, zur Erfassung des Ressourcenverbrauchs sinnvoll sein. Aufgrund ihrer Bedeutung für die Modellergebnisse sollten die Quellen, die zur Ermittlung des Ressourcenverbrauchs herangezogen wurden, in der Modelldokumentation beschrieben und gut begründet werden.

Die Schritte für die Umwandlung der Daten aus der Nutzenbewertung in eine für die Einbindung in das Modell geeignete Form sollten ebenfalls beschrieben und jede Anpassung transparent gemacht werden.

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 patientenrelevanten Endpunkte. In solchen Fällen können valide Surrogatparameter für die patientenrelevanten Endpunkte wie Langzeitmorbidity und Mortalität untersucht werden [8].

Die Interpretation jeder prädiktiven Funktion und die Validität der Übertragung feststehender Risikofunktionen auf die modellierte Population sowie vorgenommene Anpassungen müssen erklärt und begründet werden. Gründe für die Verwendung von Daten aus Beobachtungsstudien 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 Annahme, 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 des Ressourcenverbrauchs 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 Definition der Population klinische Befunde, Laborergebnisse, über Alter und Geschlecht hinausgehende Patientencharakteristika oder sonstige Parameter einschließt, die üblicherweise nicht in solchen Datenbanken codiert werden. Auch Abrechnungsdaten von Arzneimitteln können Informationen zu in der Realität üblichen Behandlungspfaden, zum Wechselverhalten, zur Adhärenz (zumindest aus Daten zur Arzneimittelbeschaffung, wenn auch nicht über die tatsächliche Arzneimittelaufnahme) und Persistenz liefern und daher im Rahmen der Kostenermittlung genutzt werden.

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 Verzögerungen in der Veröffentlichung von Daten. Obwohl eine vorhandene Datenbank nützlich sein kann, den Ressourcenverbrauch bei bestimmten Krankheitszuständen zu quantifizieren, ist bei auf Registerdaten beruhenden komparativen Analysen Vorsicht geboten, da es durch die fehlende Randomisierung der Patienten zu erheblichen Verzerrungen kommen kann [20].

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, bilden sie die aktuelle klinische Praxis gut ab.

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

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) [38]. Die Analysen, die durchgeführt wurden, um die für das 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 Umfragedaten kompiliert wurden. Diese Quellen können relevante Daten zur Demografie (z. B. Alter, Geschlecht), zu gesundheitsrelevantem Verhalten sowie Risikofaktoren (z. B. Gewicht, Raucherstatus) liefern [39]. Kompilierte Statistiken werden häufig genutzt, um Mortalität aus anderen Gründen als den für das Modell relevanten abzuschätzen (z. B. die Human Mortality Datenbank [40]). Jede im Modell verwendete Quelle kompilierter Daten muss genannt und auf Qualität und Relevanz hin bewertet werden, um sicherzustellen, dass diese angemessen, ausreichend aktuell und vollständig ist.

8.2.6 Expertenmeinung

Expertenmeinungen, Fokusgruppen und Konsensuspanels können in sehr begrenzter Weise in Modellierungsstudien verwendet werden: zur Aufstellung der Modellstruktur und zur Unterstützung der in Einflussdiagrammen und im Modellkonzept getroffenen Annahmen. Sofern keine Daten verfügbar sind, wird in gesundheitsökonomischen Studien gelegentlich ein pragmatischer Ansatz gewählt und Evidenz durch Expertenmeinungen ersetzt [13,21,41]. Dies kann jedoch als methodischer Fehler gewertet werden [42]. Für tatsächliche Inputwerte und funktionale Beziehungen ist diese Informationsquelle nahezu nie angemessen [20,43-45].

Expertenmeinungen können als wertvoll erachtet werden, wenn der gesamte Prozess des Patientenmanagements erfasst werden soll. Experten können auch den Transfer von Daten zum Ressourcenverbrauch zwischen Ländern erleichtern, wenn entsprechende Daten im eigenen Land nicht verfügbar sind. Die Rechtfertigung für die Nutzung von Expertenmeinungen und eine Beschreibung, wie diese erhoben wurden, müssen im Modellbericht aufgeführt werden [43]. Die Variabilität in den erhobenen Meinungen muss wiedergegeben werden [13,43,45,46].

9 Funktionale Beziehungen

9.1 Definition

Eine funktionale Beziehung ist eine Gleichung, die eine Responsevariable ihren Determinanten, d. h. den Variablen zuordnet, die mit der Responsevariablen korrelieren oder assoziiert sind. Unter den Kovariablen (Determinanten) 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 zutreffen. Dieses Kapitel beschreibt verschiedene Arten funktionaler Beziehungen 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 Parameter angemessen verknüpft sind, sodass eine Änderung in einem Parameter eine Änderung in einem assoziierten 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 Effektes (also der Responsevariablen) über mehrere Stufen oder Werte der ausgelassenen Eingangsvariablen. Der Einschluss von Variablen, die mit Nutzen- und Kostenparametern korreliert sind, erhöht die Modellgenauigkeit und reduziert die Unsicherheit der Ergebnisse ökonomischer Schätzungen [47-49].

9.3 Allgemeine Konzepte

Funktionale Beziehungen sind besonders wichtig bei simulationsbasierten ökonomischen Evaluationen, da diese die Simulation bestimmen. Multiple Regressionstechniken und andere Methoden wie beispielsweise Überlebenszeitanalysen können angewandt werden, um die Parameter dieser funktionalen Beziehungen zu schätzen [50-55].

Funktionale Beziehungen sollten bestimmt werden, um Nutzen- 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 der Responsevariablen assoziiert sind, einbezogen, erlaubt dies patientenspezifische Vorhersagen bei Änderung in den betroffenen Parametern. Dieses System erlaubt auch die Integration von Subgruppeneffekten durch die Einbeziehung von

Interaktionstermen. Der Ressourcenverbrauch kann in ähnlicher Weise modelliert werden, um die Abhängigkeit vom Patiententyp zu erfassen. Beispielsweise können ältere und schwer erkrankte Patienten höhere Kosten verursachen. Die statistischen Eigenschaften der Daten zum Ressourcenverbrauch (rechtsschiefe Verteilung der Kostendaten) erfordern jedoch besondere Vorsicht und Sorgfalt bei der Modellierung [56].

Wenn die einfließenden Variablen miteinander assoziiert sind, sollte der Zusammenhang zwischen den Variablen in das Modell integriert werden. Beispielsweise kann ein großer Rückgang bei den Krankenhausaufenthalten mit einer starken Zunahme des Einsatzes ambulanter Pflegedienste einhergehen und vice versa. Ebenso korrelieren Risikofaktoren häufig miteinander; beispielsweise das Gewicht mit dem Geschlecht des Patienten oder die Schwere einer Erkrankung mit den Patientencharakteristika. Um diese Zusammenhänge zu modellieren, könnten beispielsweise die Cholesky-Zerlegungsmethode (unter der Voraussetzung eines linearen Zusammenhangs) oder Bootstrap-Verfahren angewendet werden, die die Zusammenhänge zwischen den Variablen berücksichtigen [50].

Kontinuierliche Variablen können meist mit linearen Regressionstechniken modelliert werden. Dies gilt jedoch z. B. nicht immer für Kostendaten, welche zu einer (sehr) schiefen Verteilung neigen. Logarithmische Transformationen oder alternative Modelle, basierend auf Gammaverteilungen, sollten hier in Erwägung gezogen werden [47,57,58]. Wenn die Schiefe in den Kosten durch eine übermäßige Zahl von Patienten ohne Kosten verursacht wird, ist ein zweistufiges Vorgehen angebracht. Bei diesem sollte zuerst modelliert werden, mit welcher Wahrscheinlichkeit überhaupt Kosten verursacht werden. Im zweiten Schritt sollte eine Analyse der Höhe der Kosten für die Patienten erfolgen, für die Kosten anfallen [59].

Wenn Daten an mehreren Standorten erhoben werden, ist es wichtig, den Einfluss des Standorts auf die Effekte und Kostenparameter zu untersuchen [47,58,60]. Potenzielle Verzerrung muss durch die Wahl einer geeigneten Technik kontrolliert werden [58,61-63].

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 basierend auf früheren klinischen Erkenntnissen über die zu untersuchenden Beziehungen. Das IQWiG empfiehlt in Übereinstimmung mit seinen Allgemeinen Methoden die Rückwärtsselektion [18]. 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. Beispielsweise kann im Falle einer linearen Gleichung die Annahme zugrunde liegen, dass die Kriterien die erklärte Varianz maximieren sollten. Bei logistischen und Überlebenszeitanalysen können prädiktive Funktionen hinsichtlich ihrer Diskriminationskraft evaluiert werden. Maße wie der c-Index oder Receiver Operating Characteristic-(ROC)-Kurven können verwendet werden [51].

Die der Erstellung der funktionalen Beziehungen zugrunde liegenden 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 [51,64]. Wenn kausale Funktionen entwickelt werden, sollten Annahmen zur Existenz und Richtung kausaler Beziehungen anhand vom Kausalgraphen dargestellt werden [65].

10 Modellierungstechniken

Verschiedene Modellierungstechniken sind verfügbar und wurden bereits in gesundheitsökonomischen Evaluationen eingesetzt, wie z. B. Entscheidungsbäume, die Markov-Modellierung, die diskrete Ereignissimulation, die agentenbasierte Modellierung, die Transmissionsmodellierung und andere [4,7,12,66-70].

Die Darstellung von Modellierungstechniken in den folgenden Absätzen ist keineswegs erschöpfend, aber sie beinhaltet die in der Gesundheitsökonomie am meisten verwendeten Techniken. Als weiterführende Literatur werden die Leser auf Brennans Taxonomie der Modellstrukturen [7], Koopmans Untersuchung von Infektionserkrankungsmodellen [71] und Stahls Übersicht zu Modellierungsmethoden [69] verwiesen.

10.1 Die Wahl der Modellierungstechnik

Es ist 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, der entsprechenden Erkrankung und den Rahmenbedingungen. Der geeignete Prozess zur Entwicklung eines Modells, das alle relevanten Bewertungsprobleme berücksichtigt, erfordert vorab, das Modellkonzept vollständig zu konstruieren und zu verstehen, die verfügbaren Daten zu erheben und die funktionalen Beziehungen zu erfassen. Erst dann sollte eine Modellierungstechnik gewählt werden. Die vom IQWiG im Einzelfall eingesetzte Modellierungstechnik wird projektspezifisch im vorläufigen Berichtsplan beschrieben und zur Diskussion gestellt.

Letztlich 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 es auch in der Modellierung Unkundigen zu erleichtern, ein Modell nachzuvollziehen, sind nicht zielführend. Genauso sollte die Wahl der Modellierungstechnik nicht aus Gewohnheit erfolgen. Mit einer adäquaten Dokumentation können selbst komplexe Modelle transparent dargestellt werden. Zunehmende Rechnerkapazitäten und neuere Techniken der Varianzreduktion führen dazu, dass die Zeit und die technischen Anforderungen zur Berechnung komplexer Modelle beträchtlich abnehmen [9,70,72,73]. Führt die Wahl der Modellierungstechnik zu einer Modifikation des Modellkonzeptes, ist dies ein Hinweis darauf, dass die gewählte Modellierungstechnik falsch sein könnte [5,74].

10.2 Aspekte bei der Wahl der Modellierungstechnik

Obgleich viele Modellcharakteristika bei der Wahl der Modellierungstechnik berücksichtigt werden könnten [7], sind zwei Aspekte für gesundheitsökonomische Evaluationen besonders wichtig. Der erste Aspekt ist die Ebene, auf welcher die Population modelliert wird (Kohorten- versus Individualebene). Der zweite Aspekt ist die Form, in der Krankheitsabläufe dargestellt werden (zustandsübergangsbezogene Simulation versus Ereignissimulation).

10.2.1 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 Kohortenmodellen 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, 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 Populationsniveau darstellen. Für jedes Individuum wird aufgrund seiner einzigartigen Charakteristika ein individueller Verlauf modelliert. Obgleich diese individuellen Modelle mehr Daten, Rechnerkapazität und Laufzeiten [70,72] benötigen, sind sie nicht notwendigerweise schwieriger mit Daten zu füllen und zu verstehen.

Kohortenmodelle sind die heutzutage in der gesundheitsökonomischen Evaluation am meisten verwendeten Techniken [12,75]. Die individuelle Simulation aber (häufig auch als Mikrosimulation bezeichnet), die früher in der gesundheitsökonomischen Evaluation kaum bekannt war, hat in den zurückliegenden Jahren beträchtlich an Bedeutung gewonnen [70,72]. Auch wenn es Widerstände gegen ihre Verwendung gibt [76], wird doch weithin akzeptiert, dass eine Simulation auf Individualebene 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 Patientencharakteristika und ihre Heterogenität in der Zielpopulation keine große Relevanz für das Entscheidungsproblem haben, kann die Kohortensimulation ausreichend sein, um die wissenschaftliche Fragestellung zu beantworten [4,7,9,50,68,70,72,77,78].

Für die gesundheitsökonomischen Evaluationen 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 zu diesen Verteilungen keine unrealistischen Vereinfachungen der Wirklichkeit darstellen.

10.2.2 Ereignisbasierte versus zustandsübergangsbasierte Simulation

Die für Modellierungen im Allgemeinen am häufigsten verwendete Technik betrachtet die Ereignisse, die eintreten können, im Rahmen einer diskreten Ereignissimulation [79] oder verwandter Techniken [80]. Bei gesundheitsökonomischen Evaluationen ist jedoch das sogenannte Zustandsübergangskonzept (Markov-Modell) weitaus verbreiteter [75]. Markov-Modelle wurden in der gesundheitsökonomischen Evaluation in den frühen 1980er-Jahren eingeführt [81] und blicken auf eine lange Geschichte im Bereich der Entscheidungsfindung im Gesundheitswesen zurück [82].

In einem Markov-Modell durchlaufen die Patienten unterschiedliche, sich gegenseitig ausschließende Gesundheitszustände (Markov-Zustände) in Zeitintervallen (Markov-Zyklen). Während dieses Prozesses werden kumulative Gesundheitseffekte und Kosten berechnet. Diese können dann zwischen den betrachteten Interventionen verglichen werden [83]. Die Technik der Markov-Modellierung einschließlich ihrer wichtigsten Vor- und Nachteile wird in Abschnitt 10.3.3 ausführlich beschrieben.

Im Gegensatz zu den zustandsübergangsbasierten Modellen konzeptionalisiert die ereignisbasierte Simulation [79] die Realität im Sinne der Ereignisse, die auftreten können. Diese Form der Simulation erlaubt eine Abbildung des natürlichen Krankheitsverlaufes, der Behandlungsmuster, der Adhärenz und anderer relevanter Faktoren. 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. Die diskrete Ereignissimulation wird in Abschnitt 10.3.2 ausführlich beschrieben.

Ereignisbasierte Modelle sind sehr leistungsstarke Instrumente, die gegenwärtig in der Gesundheitsökonomie noch zu selten Anwendung finden. Ausgehend von der Flexibilität und der Realitätsnähe ereignisbasierter Simulationen wird empfohlen, dass diese Modelle in derselben Weise für IQWiG-Bewertungen in Betracht gezogen 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 erlaubt [72].

Die Wahl eines zustandsübergangs- anstelle eines ereignisbasierten Modellierungsansatzes muss begründet werden. Auch sollte dargelegt werden, dass die Wahl einer bestimmten Modellierungstechnik zu keinen unangemessenen Veränderungen des Modellkonzepts oder unakzeptablen Realitätsverzerrungen führt.

10.3 Techniken

10.3.1 Entscheidungsbaumanalysen

Ein Entscheidungsbaum ist eine visuelle Darstellung aller möglichen Optionen und Konsequenzen, die sich als Folge einer untersuchten Technologie ergeben können [84]. 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 den Patientencharakteristika abhängen. Am Ende des Entscheidungsbaums führt jeder Pfad zu einem Outcome, z. B. den alternativen Endpunkten einer klinischen Studie. 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 einfache Entscheidungsprobleme, bei denen die Wiederholung relevanter Ereignisse limitiert ist und ein kurzer, feststehender Zeithorizont analysiert wird. Entscheidungsbäume können entweder mittels Kohortensimulation oder Simulation auf Individualebene berechnet werden.

10.3.2 Diskrete Ereignissimulation

Auf dem Gebiet der Gesundheitsökonomie hat der Einsatz der diskreten Ereignissimulation in den letzten Jahren an Bedeutung gewonnen [4,67,69,72,77,85]. Die Grundlagen der diskreten Ereignissimulation sind allerdings in der nicht medizinischen Literatur am besten dokumentiert [86-94]. In diesem Abschnitt wird diese Technik aufgrund ihrer steigenden Bedeutung bei gesundheitsökonomischen Evaluationen umfassend dargestellt und ihre Vor- und Nachteile im Vergleich zu den zustandsübergangsbasierten Techniken werden hervorgehoben.

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 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 „Enterereignis“ gebildet und einem Typ zugeordnet. In gesundheitsökonomischen Evaluationen ist es üblich, sämtliche Entitäten bereits zu Beginn der Simulation zu kreieren, was jedoch nicht unbedingt erforderlich ist. 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 „Endereignis“ auftritt 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.

Entitäten zeichnen sich dadurch aus, dass ihren Attributen spezielle Werte zugeordnet werden. Alle Entitäten teilen einen gemeinsamen Satz von Attributen (z. B. hätten alle das Attribut Alter, 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). Veränderungen können auch durch andere Effekte, die das gesamte System betreffen (z. B. Veränderungen in der Erstattung, Einführung eines neuen Prozesses), oder durch Geschehnisse, die andere Entitäten betreffen (z. B. Tod einer Pflegekraft), verursacht werden. 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, 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 Halbzykluskorrekturen oder die Anwendung anderer Maßnahmen zur Reduktion der durch die Vorgabe fester Zeitzyklen bedingten Abweichungen. Dem Analytiker wird ermöglicht, alle bestehenden Zeitabhängigkeiten sachgemäß umzusetzen und das zeitliche Auftreten der Ereignisse sowie die Dauer der Krankheitszustände etc. zu verfolgen. Verzögerungen – sowohl geplante (z. B. ein Tag im Krankenhaus wegen eines Eingriffs) als auch fremdbestimmte (z. B. Warten auf ein verfügbares Bett) – können integriert werden.

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. ärztliches Personal, Pflegepersonal, weitere medizinische Berufsgruppen), Güter (Medikamente, Katheter), Organe (z. B. Lebertransplantate), Platz (z. B. Krankenhausbetten) oder andere Dinge sein, die als Ressource verstanden werden. Den Ressourcen sind Kosten pro Einheit zugeordnet, welche durch den Gebrauch pro Stunde usw. erfasst werden und die zeitabhängig sein können (z. B. teurer außerhalb der normalen Arbeitszeit); selbst die Kosten von Inaktivitätszeiten können abgebildet 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 die benötigte 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.

Ebenfalls ein Vorteil der diskreten Ereignissimulation ist es, dass Individuen miteinander interagieren können [4,7,89]. Beispielsweise erlaubt es die diskrete Ereignissimulation, in ein Modell zu integrieren, dass individuelle Entitäten um beschränkte Ressourcen konkurrieren und sich Warteschlangen bilden können. Im Gesundheitswesen kann die Bewertung, in welcher Weise neue Interventionen entlastend auf limitierte Ressourcen wirken, bedeutsam sein. Ein anderes Beispiel sind Verhaltensänderungen. So 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 infektiöser Erkrankungen, bei denen die Übertragung von Infektionen eine unmittelbare Funktion des Infektionsstatus der Individuen in der Gemeinschaft darstellt, hilfreich, die Interaktionen und Ansteckungswege zwischen den Entitäten abzubilden und einzubeziehen (siehe Abschnitt 10.3.5).

Diskrete Ereignissimulationen und auf der Individualebene analysierte Markov-Modelle (Abschnitt 10.3.3) benutzen verschiedene Arten von Funktionen, die mathematisch erfasst werden. Eine diskrete Ereignissimulation verwendet auch Verteilungen wie Exponential-, Weibull-, Beta-, Gammaverteilungen u. a., um die Variabilität der Inputparameter und Ereigniszeiten etc. abzubilden. Zusammen erzeugen diese das stochastische Verhalten des Modells. Zufallszahlen müssen mittels hoch entwickelter 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 ihrer Erfahrungen, einzigartig. Dies ergibt einen unterschiedlichen Satz von Resultaten für jede einzelne Replikation. Nach Durchführung vieler Replikationen werden sich die Ergebniswerte ihrem Mittelwert annähern und ergeben so eine stabilere Schätzung der Resultate. Häufig wird die Frage gestellt, wie viele Replikationen notwendig seien. Auf diese Frage gibt es keine allgemeingültige Antwort, weil dies von der Variabilität innerhalb des Modells und zwischen den Patienten etc. abhängig ist. Am besten lässt sich dies empirisch überprüfen, indem Standardabweichungen der kritischen Ergebnisse zwischen den Replikationen untersucht werden. 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 Behandlungseffektes anstatt aus der Verteilung der Behandlungseffekte gezogene Werte). 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.

Moderne Software ermöglicht eine sehr transparente Darstellung der Beziehungen innerhalb der diskreten Ereignissimulation, z. B. der Patientenströme innerhalb der Simulation. Softwareanwendungen wie Arena[®] [89] erlauben die Programmierung 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.

10.3.3 Markov-Modelle

Markov-Modelle sind besonders geeignet 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), (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, zweiter Myokardinfarkt). Es wurden bereits zahlreiche Anleitungen zur Konstruktion von Markov-Modellen publiziert [75,82,95]. In diesem Abschnitt werden die Technik sowie einige grundlegende Stärken und Schwächen von Markov-Modellen beschrieben. Die Ausweitung dieser Modelltechnik auf eine Simulation auf Individualebene wird beschrieben.

In einem zustandsübergangsbasierten Markov-Modell startet die gesamte Kohorte zum Zeitpunkt null entweder vom gleichen Initialzustand oder entsprechend einer vorgegebenen Initialverteilung von verschiedenen Markov-Zuständen. In jedem Zustand wird die Population 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 [72,75]. 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 können Auswirkungen auf die Bewertung des Zustandes haben (z. B. hängen die kumulierten Kosten während eines Zyklus von der Krankheitsschwere ab)

[95]. Die Charakteristika jedes Zustands können sich auch im Laufe der Zeit ändern. 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, welches aufzeigt, wie viele Patienten sich im Laufe der Zeit in dem jeweiligen Zustand befinden [95,96].

Zustandsübergangsmodele 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 in einem bestimmten Zyklus befinden [82]. Diese diskreten Zustände schließen sich gegenseitig aus (kein Teil der Population kann in demselben Zyklus zwei Zuständen angehören), werden als homogen angenommen und sind vollständig (die Summe der Kohortenanteile in allen Zuständen zusammengenommen muss in jedem Zyklus 100 % entsprechen). Aus absorbierenden Zuständen (z. B. Tod, Amputationen) gibt es keinen Austritt: Der Teil der Kohorte, der diesen Zustand erreicht, verbleibt dort für den Rest der Modellierungsdauer. 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 wird die Zeit in Markov-Modellen 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 [72]. Der Analyst muss eine ausreichend kleine Zykluslänge wählen, um den kontinuierlichen Zeitfluss in der Wirklichkeit nachzuahmen.

Eine wichtige Annahme bei Markov-Modellen ist die Annahme einer „Gedächtnislosigkeit“ [70,75,95]. Dies bedeutet, dass Personen sich von einem Zustand in den anderen bewegen und hierbei alle Charakteristika des neuen Zustandes erwerben und alle Charakteristika zurückliegender Zustände verlieren. Die Wahrscheinlichkeit des Wechsels von einem Zustand in den anderen hängt einzig und allein vom gegenwärtigen Gesundheitszustand ab. Diese Wahrscheinlichkeit ist weder abhängig von vorhergehenden Erfahrungen noch von der Aufenthaltsdauer in einem bestimmten Zustand. Diese Annahme der „Gedächtnislosigkeit“ spiegelt jedoch nicht die Realität, denn oftmals wird die Übergangswahrscheinlichkeit zwischen verschiedenen Krankheitsstadien von Faktoren bestimmt, die aus früheren Stadien resultieren. Das bedeutet, dass relevante Erfahrungen wie vorhergehende Ereignisse in die Zustandsbeschreibungen einbezogen werden müssen. Hierzu können sogenannte „Tunnelzustände“ erstellt werden, die die Zeit abbilden, die ein Patient bereits in einem spezifischen Krankheitsstadium zugebracht hat. Dies kann zu einer großen Anzahl von Zuständen führen und dadurch das Modell unhandlich machen.

Es kann ebenfalls vorkommen, dass die Charakteristika der Zustände nicht alle wichtigen Eigenschaften umfassen, die für eine Modellierung der Übergänge erforderlich sind, was wiederum zu Verzerrungen führen kann (z. B. in Einjahreszyklen wird das mittlere Alter

üblicherweise um ein Jahr erhöht, obwohl die älteren Personen aufgrund von Krankheit oder Tod ausscheiden können). Auch dieser Aspekt kann durch die Einführung zusätzlicher, hinreichend homogener Gesundheitszustände überwunden werden. Bei multiplen Patientencharakteristika kann dies allerdings schnell zu einem expandierenden Satz an Zuständen führen und eine Implementierung gefährden. In beiden zuvor beschriebenen Fällen kann eine Simulation auf der Individualebene mittels Monte-Carlo-Simulation sinnvoll sein. Diese wird am Ende des Abschnitts ausführlicher beschrieben.

Die einfachsten Markov-Modelle sind Ketten, welche der Annahme unterliegen, dass die Übergangswahrscheinlichkeiten über die gesamte Modellierungsdauer hinweg konstant sind [82]. Markov-Ketten sind eine angenehme Art zu modellieren, da sie numerisch berechnet werden können [97], 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. Beispielsweise ändert sich die Hintergrundmortalitätsrate von Individuen mit zunehmendem Alter. Da entscheidungsanalytische Software iterative Berechnungen in Markov-Modellen erlaubt, die nicht der technischen Komplexität der Matrixalgebra 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) [81]. Ein weiterer Aspekt von Markov-Modellen in Bezug auf Übergangswahrscheinlichkeiten ist die bereits beschriebene Markov-Annahme der „Gedächtnislosigkeit“, die dazu führt, dass eine ausreichend große Menge an Zuständen gewählt werden muss, um die Patientengeschichte abzubilden und ein realistisches Modell der Krankheit zu erstellen.

In einem reinen Markov-Modell werden Interaktionen zwischen Individuen nicht betrachtet [4,7,67]. 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. Bei dieser Art der Modellierung durchläuft anstelle einer Kohorte eine große Anzahl einzelner Patienten das Markov-Modell. Dies ermöglicht es, die Krankheitsgeschichte jedes Individuums zu berücksichtigen. Die Monte-Carlo-Simulation überwindet damit die restriktive (und häufig unrealistische) Markov-Annahme einer „Gedächtnislosigkeit“ im Modell und ermöglicht eine Modellierung unter Berücksichtigung der gesamten Krankengeschichte und der Variabilität der Charakteristika der eingehenden Patienten [82]. Hierdurch können viele der Limitationen, die den Kohortenmodellen eigen sind, umgangen werden [7,75,82]. 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 von Gleichungen nutzen, um Modellierungsprozesse zu kontrollieren. Markov-Modelle können überdies in unterschiedlicher Weise modifiziert werden, um die Limitationen eines reinen Markov-Modells zu umgehen. Beispielsweise werden „Tunnelzustände“, „Halbzykluskorrekturen“, 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 Instrumente, die die Visualisierung der Zustandsverteilungen im Zeitablauf (die sogenannte Markov-Spur), die Durchführung von Sensitivitätsanalysen, die Darstellung von Unsicherheiten sowie die Anbindung an Tabellenkalkulationsprogramme und zur Makrosprache erleichtern bzw. unterstützen.

10.3.4 Agentenbasierte Simulation

Die agentenbasierte Simulation ist eine Weiterentwicklung der diskreten Ereignissimulation, wird aber seltener als Modellierungstechnik im Gesundheitswesen verwendet [66,80,98-100]. 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 [80]. Diese Technik ist gut 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 hat. Obwohl die einfache diskrete Ereignissimulation viele dieser Anforderungen bewältigen kann, hat die agentenbasierte Simulation den Vorteil, dass sie Interaktionen berücksichtigt. 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 soziale Einheiten 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“ [80]. 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. Darüber hinaus ermöglicht diese Software eine unmittelbare visuelle Darstellung der Modellprozesse und erhöht dadurch die Transparenz der Modelle [66]. Diese visuellen Hilfsmittel sind nicht nur wertvoll für den Entscheidungsträger, sondern auch für den Analysten bei der Fehlersuche.

Einige Szenarien, bei denen agentenbasierte Simulation in Betracht kommen [66], 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.
- Netzwerkeffekte können vorhanden sein.
- In dem System sind Schwankungen möglich, wobei ein stabiler Zustand, basierend auf durchschnittlichen Verhältnissen, unrealistisch ist.

10.3.5 Transmissionsmodelle

Bei der Modellierung von Infektionskrankheiten in den Bereichen Epidemiologie und Public Health stellen Transmissionsmodelle eine weitverbreitete und anerkannte Technik dar [71]. Transmissionsmodelle [101] können mit verschiedenen Komplexitätsgraden konstruiert werden und können entweder deterministisch oder stochastisch sein [68,71]. 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. Beispielsweise 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 [68,71]. 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 [71]. Diese Annahmen sind häufig nicht haltbar. Einige der Annahmen in Transmissionsmodellen können bei anspruchsvollerer Modellierung der Krankheitsübertragung gelockert werden. Techniken, die speziell für solche Bereiche entwickelt wurden, bieten geeignete und präzise Modellierungsansätze [71].

11 Umgang mit Variabilität und Unsicherheit

11.1 Arten von Variabilität

In ökonomischen Modellen kommen vornehmlich zwei Arten von Variabilität vor [84]. Es gibt eine Variabilität, die sich per se nicht vermeiden lässt und die durch Unterschiede in der Ausprägung von Variablen bei unterschiedlichen Patienten entsteht. Eine zweite Form der Variabilität resultiert aus der Unvollständigkeit der Informationen, die einem Modell zugrunde liegen, d. h. aus den Annahmen, auf denen dieses basiert. Die erste Form der Variabilität hängt von den Charakteristika der Intervention, den Patienten, dem Kontext und den Stückkosten ab. Entweder sind die Werte dieser Variablen mit Sicherheit bekannt (z. B. Alter, Geschlecht) oder aber es besteht Unsicherheit über ihren Wert (z. B. Tageskosten eines Krankenhausaufenthalts, Effektschätzer aus klinischen Studien). Im ersten Fall spricht man eher von Heterogenität als von Variabilität. Die zweite Form der Variabilität ist der Tatsache geschuldet, dass Modelle die Realität aufgrund ihrer Komplexität nicht vollkommen erfassen und wiedergeben können und somit simplifizierende Annahmen treffen müssen, um überhaupt kalkulierbar, nachvollziehbar und transparent zu bleiben. Diese Art der Variabilität ist modellinhärent. Für beide Arten der Variabilität gibt es methodische Ansätze, um mit ihnen zielführend umzugehen.

11.2 Unsicherheit und deren Quantifizierung

Modelle werden mit Daten aus verschiedenen Quellen gespeist, z. B. klinischen Studien, Beobachtungsstudien, Datenbankauswertungen, Befragungen usw. Somit fließen in Modelle Daten ein, die mit Unsicherheit behaftet sind [102]. Die Ursachen dieser Unsicherheit lassen sich auf mehrere Arten beschreiben [37,50,103-105].

(1) Die Unsicherheit vieler Modellparameter ergibt sich daraus, dass sie Variabilität unterliegen, weil ihr Wert aus Stichproben mit endlicher Stichprobengröße geschätzt wird. Die Unsicherheit der Parameter spiegelt unser Unvermögen wider, diese Werte genau zu bestimmen, beispielsweise die Wirksamkeit einer Intervention oder deren Ressourcenverbrauch [106]. Diese Art von Unsicherheit wird oft durch Konfidenzintervalle oder andere statistische Ansätze zur Beschreibung von Varianz erfasst, welche die beobachtete Unsicherheit des jeweiligen Parameters in der Datenquelle widerspiegeln.

(2) Bei den Kosten kann Unsicherheit über ihre Determinanten bestehen, z. B. Dosierung eines Arzneimittels über die Zeit. Auch kann das Modell stochastisch ausgelegt sein (es verwendet Zufallszahlen bei den Monte-Carlo-Ziehungen). Um diese Art von Unsicherheit einzugrenzen, werden verschiedene Techniken angewendet. Diese Techniken zur Varianzreduktion (z. B. Klonen) werden in der Literatur ausführlich beschrieben [73,107,108]. Eine verbreitete Technik ist die wiederholte Stichprobenziehung (Bootstrapping). Die nach Anwendung all dieser Techniken verbleibende Varianz muss quantifiziert werden.

(3) Die Heterogenität der Patientenpopulation wie Alter, Geschlecht, relevante klinische Vorgeschichte wirkt sich darauf aus, wie die Ergebnisse der ökonomischen Evaluation zu verwenden sind. Obwohl Patientencharakteristika in einer Population heterogen sind, impliziert dies nicht per se Unsicherheit [103]. Indes kann Unsicherheit aufgrund unvollständiger Information über diese Charakteristika vorliegen. Auch das Ausmaß dieser Unsicherheit sollte zusammen mit anderen Parameterunsicherheiten bestimmt werden.

(4) Auch aus der im vorherigen Kapitel beschriebenen Form von Variabilität in den Modellannahmen resultiert Unsicherheit, die bei der Modellierung berücksichtigt werden muss. Der Modellierer muss die Entscheidungen zur Modellstruktur und den getroffenen Annahmen offenlegen und dokumentieren. Diese können unter anderem in Zusammenhang mit der Konzeptualisierung der gesundheitlichen Intervention, den einbezogenen klinischen Ereignissen, den funktionalen Beziehungen, der Auswahl der berücksichtigten Kosten- und Nutzenkomponenten sowie den Methoden zur Messung und Bewertung der Outcomes stehen [45,102,103,109].

(5) Selbst zuvor festgesetzte Inputparameter wie die Diskontierungsrate können sich ebenfalls in einer Spanne von Werten bewegen und damit variieren.

11.3 Umgang mit Unsicherheit und ihre Auswirkungen auf die Ergebnisse

Parameterunsicherheit sowie andere Unsicherheiten, die nicht reduzierbar sind, sollten quantifiziert werden. Alle dazu unternommenen Analysen sollten hinsichtlich der Spannweiten für die eingesetzten Parameterwerte und getroffenen Annahmen vollständig dokumentiert werden.

Unsicherheiten bei den Modellparametern und -annahmen müssen im Modell berücksichtigt werden und ihr Einfluss auf die modellierten Ergebnisse unter Verwendung geeigneter Methoden quantifiziert werden [26,103,110]. Die drei gebräuchlichsten Methoden zur Überprüfung der Robustheit der Modellergebnisse in Anbetracht von Unsicherheit sind in gesundheitsökonomischen Evaluationen die univariate deterministische, die multivariate deterministische und die multivariate probabilistische Sensitivitätsanalyse. Weitere Methoden [107,111] werden selten in Health Technology Assessments angewandt. Multivariate Analysen sind insbesondere dann von Vorteil, wenn die Analyse anhand spezifischer, durch einen Satz von Variablenwerten definierter Szenarien erfolgt.

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 Werten gehalten werden. Hierbei ist zu beachten, dass Korrelationen zwischen Inputvariablen nicht berücksichtigt werden könnten. Die univariate Sensitivitätsanalyse kann einen Anhaltspunkt zum Ausmaß des Einflusses dieser Unsicherheit geben, aber offensichtlich nicht die gesamte Unsicherheit der Ergebnisse quantifizieren. Univariate Sensitivitätsanalysen sind 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 realistischer 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 ungenau, wenn Korrelationen unter den Parameterwerten unberücksichtigt bleiben.

In der probabilistischen Sensitivitätsanalyse wird jedem Parameter, der mit Unsicherheit behaftet ist, eine Verteilung zugeordnet. Anschließend wird aus diesen Verteilungen simultan eine Stichprobe gezogen, um ein Set von Eingangswerten für einen Modelldurchlauf zu erhalten [45]. Dieser Vorgang wird viele Male wiederholt. 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.

Es wird empfohlen, univariate Sensitivitätsanalysen nicht durch multivariate oder probabilistische Sensitivitätsanalysen zu ersetzen, sondern Letztere zu ergänzen. Für univariate und multivariate Analysen sollten die Ergebnisse sowohl tabellarisch als auch in einem Tornado-Diagramm dargestellt werden, indem die entsprechenden Inputparameter-spannweiten den Ergebnisspannweiten gegenübergestellt werden. Im Falle probabilistischer Sensitivitätsanalysen sollten die Ergebnisse als kumulative Kostenverteilungen präsentiert werden. Kosten-Effektivitäts-Akzeptanzkurven [112] werden in den IQWiG-Evaluationen nicht verlangt, da bei IQWiG-Evaluationen keine Schwellenwerte von Kosten-Effektivitäts-Verhältnissen zum Einsatz kommen. Strukturelle Sensitivitätsanalysen sollten durchgeführt werden, um den Einfluss einer Variierung der strukturellen Annahmen zu eruieren.

Modellierungen, die über einen studienbelegten Zeitraum hinausgehen, sind mit einem größeren Unsicherheitspotenzial behaftet als Modellierungen, die sich auf einen studienbelegten Zeitraum beschränken. Entsprechend steigt die Bedeutung einer ausführlichen und gut dokumentierten Prüfung der Robustheit der Ergebnisse des Modells bei größeren Modellierungszeiträumen.

12 Validierung

Ansätze zur Validierung sind auf unterschiedliche Weise klassifiziert worden. Auf diese wird hier nicht vertiefend eingegangen. Im Kontext der Modellierung wird jedoch dargestellt, welche Schritte eine Validierung zwingend beinhalten sollte.

Ein Schlüsselement der Validierung befasst sich mit der Frage, ob das Modell Sinn ergibt. Die Plausibilität (Augenscheinvalidität) bezieht sich auf das Einflussdiagramm, das Modellkonzept, die Datenbeschaffung, die Erarbeitung der funktionalen Beziehungen und die Auswahl der Modellierungstechnik. Jeder Schritt muss mit dem vorhandenen Wissen, der besten verfügbaren Information und unter Berücksichtigung der „best practice“ durchgeführt werden. Ein zweites Schlüsselement der Validierung setzt sich mit der korrekten Implementierung des Modells auseinander (wird auch als technische Validierung 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 Bedeutung, weil ein technisch gutes, von Programmierfehlern freies Modell mit starker Plausibilität von geringem Wert für die Entscheidungsfindung ist, wenn es das System und sein Verhalten nicht richtig wiedergibt.

Ein Simulationsmodell, das für einen Zweck valide ist, kann für einen anderen nicht valide sein [113]. Der externe Validierungsprozess muss daher jede intendierte Verwendung des Modells abdecken und bei Einsatz für andere Zwecke muss die Validierung erneut erfolgen. Über das Vorgehen zur Validierung eines Modells herrscht Uneinigkeit, aber es gibt einige grundlegende Schritte, die eingehalten werden müssen [114].

Das Design und die Struktur des Modells sollten von Sachverständigen im Fachgebiet 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. Hierdurch sollen ihre korrekte Wiedergabe und die richtige Ergebniserzeugung für die eingegebenen Werte sichergestellt werden. Die Gleichungen sollten auch anhand von Extremwerten, die sich an den Modellrändern orientieren, berechnet werden, um zu überprüfen, ob 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). 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 historischen Situationen wie beispielsweise bereits durchgeführten klinischen Studien oder Registern stammen, aus welchen bislang keine Daten in das getestete Modell eingeflossen sind. Die Modellergebnisse werden dann mit den bereits bekannten Ergebnissen verglichen [9]. Durch Abgleich mit 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 Abweichungen von der Realität darstellen können. 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 [6,12,24,26,110] 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 entspricht
- Überblick über die eingesetzten Validierungstechniken und ihre Ergebnisse
- Detaillierte Ergebnisdarstellung einschließlich einer Bewertung des Einflusses von:
 - Anwendung der Gesundheitstechnologie in relevanten Subgruppen
 - 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 strukturellen 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. Siebert U. Entscheidungsanalytische Modelle zur Sicherung der Übertragbarkeit internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems: ein Methodenbeitrag zu HTA [online]. 2005 [Zugriff: 09.10.2009]. (Schriftenreihe Health Technology Assessment; Band 16). URL: http://gripsdb.dimdi.de/de/hta/hta_berichte/hta099_bericht_de.pdf.
4. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004; 9(2): 110-118.
5. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T et al. Modelling in economic evaluation: an unavoidable fact of life. *Health Econ* 1997; 6(3): 217-227.
6. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C et al. Principles of good practice for decision analytic modeling in health care evaluation: report of the ISPOR task force on good research practices-modeling studies. *Value Health* 2003; 6(1): 9-17.
7. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; 15(12): 1295-1310.
8. Siebert U. When should decision-analytic modeling be used in the economic evaluation of health care? *Eur J Health Econ* 2003; 4(3): 143-150.
9. Eddy DM. Accuracy versus transparency in pharmacoeconomic modelling. *Pharmacoeconomics* 2006; 24(9): 837-844.
10. Ades AE, Claxton K, Sculpher MJ. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. *Health Econ* 2006; 15(4): 373-381.
11. Chapman GB, Sonnenberg FA. Decision making in health care: theory, psychology, and applications. Cambridge: Cambridge University Press; 2000.
12. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.
13. Philips Z, Bojke L, Sculpher MJ, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006; 24(4): 355-371.
14. Eddy DM. Technology assessment: the role of mathematical modeling. Washington: National Academy Press; 1985.

15. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden zur Bewertung von Verhältnissen zwischen Nutzen und Kosten: Version 1.0. Köln: IQWiG; 2009.
16. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Arbeitspapier Kostenbestimmung: Version 1.0. Köln: IQWiG; 2009.
17. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2000.
18. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Methoden: Version 3.0 [online]. 27.05.2008 [Zugriff: 09.10.2009]. URL: http://www.iqwig.de/download/IQWiG_Methoden_Version_3_0.pdf.
19. Leidl R, Graf von der Schulenburg JM, Wasem J (Ed). Ansätze und Methoden der ökonomischen Evaluation: eine internationale Perspektive. Baden-Baden: Nomos; 1999. (Schriftenreihe Health Technology Assessment).
20. Nuijten MJC. The selection of data sources for use in modelling studies. *Pharmacoeconomics* 1998; 13(3): 305-316.
21. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application. *Pharmacoeconomics* 2000; 17(5): 461-477.
22. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Economic evaluation using patient level data. In: *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2005. S. 247-255.
23. Ramsey SD, Willke R, Briggs A, Brown R, Buxton M, Chawla A et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health* 2005; 8(5): 521-533.
24. Soto J. Health economic evaluations using decision analytic modeling. *Int J Technol Assess Health Care* 2002; 18(1): 94-111.
25. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-149.
26. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: NICE; 2004. URL: http://www.nice.org.uk/niceMedia/pdf/TAP_Methods.pdf.
27. Song F, Altman DG, Glenny A, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published metaanalyses. *BMJ* 2003; 326(7387): 472.
28. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc* 2006; 101(474): 447-459.

29. Lu G, Ades AE, Sutton AJ, Cooper NJ, Briggs AH, Caldwell DM. Meta-analysis of mixed treatment comparisons at multiple follow-up times. *Stat Med* 2007; 26(20): 3681-3699.
30. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.
31. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002; 21(16): 2313-2324.
32. Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.
33. Nixon RM, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. *Stat Med* 2007; 26(6): 1237-1254.
34. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008; 26(9): 753-767.
35. Sculpher MJ, Drummond MF. Analysis sans frontières: can we ever make economic evaluations generalisable across jurisdictions? *Pharmacoeconomics* 2006; 24(11): 1087-1099.
36. Barbieri M, Drummond M, Willke R, Chancellor J, Jolain B, Towse A. Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: lessons for inferring generalizability. *Value Health* 2005; 8(1): 10-22.
37. Drummond MF, McGuire AE. *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001.
38. Agency for Healthcare Research and Quality. *Registries for evaluating patient outcomes: a user's guide*; AHRQ publication no. 07-EHC001-1. Rockville: AHRQ; 2007. URL: <http://effectivehealthcare.ahrq.gov/repFiles/PatOutcomes.pdf>.
39. Statistisches Bundesamt. Statistisches Bundesamt Deutschland [online]. [Zugriff: 09.10.2009]. URL: <http://www.destatis.de>.
40. The Human Mortality Database [online]. [Zugriff: 09.10.2009]. URL: <http://www.mortality.org>.
41. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 1996; 276(15): 1253-1258.
42. Drummond MF, Sculpher MJ. Common methodological flaws in economic evaluations. *Med Care* 2005; 43(7 Suppl): 5-14.
43. Commonwealth Department of Health and Ageing. *Guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee: including major submissions involving economic analyses*. Canberra: Commonwealth of Australia; 2002. URL: [http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-general-pubs-guidelines-index.htm/\\$FILE/guidelines.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-general-pubs-guidelines-index.htm/$FILE/guidelines.pdf).

-
44. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Collection and analysis of data. In: *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2000. S. 234-235.
45. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. Oxford: Oxford University Press; 1996.
46. Ramsey SD, McIntosh M, Etzioni R, Urban N. Simulation modeling of outcomes and cost effectiveness. *Hematol Oncol Clin North Am* 2000; 14(4): 925-938.
47. Nixon RM, Thompson SG. Methods for incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. *Health Econ* 2005; 14(12): 1217-1229.
48. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. *Health Econ* 2002; 11(5): 415-430.
49. Vazquez-Polo FJ, Negrin Hernandez MA, Lopez-Valcarcel BG. Using covariates to reduce uncertainty in the economic evaluation of clinical trial data. *Health Econ* 2005; 14(6): 545-557.
50. Briggs AH, Claxton K, Sculpher M. *Decision modelling for health economic evaluation: handbooks for health economic evaluation*. Oxford: Oxford University Press; 2006.
51. Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York: Springer; 2001.
52. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004; 13(5): 461-475.
53. Kleinbaum DG, Klein M. *Survival analysis: a self-learning text*. New York: Springer; 1996.
54. Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. Hoboken: Wiley; 2002.
55. Lawless JF. *Statistical models and methods for lifetime data*. Hoboken: Wiley; 2003.
56. Nixon RM, Thompson SG. Parametric modelling of cost data in medical studies. *Stat Med* 2004; 23(8): 1311-1331.
57. Kilian R, Matschinger H, Löffler W, Roick C, Angermeyer MC. A comparison of methods to handle skew distributed cost variables in the analysis of the resource consumption in schizophrenia treatment. *J Ment Health Policy Econ* 2002; 5(1): 21-31.
58. Manca A, Rice N, Sculpher MJ, Briggs AH. Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. *Health Econ* 2005; 14(5): 471-485.

-
59. Mullahy J. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *J Health Econ* 1998; 17(3): 247-281.
60. Cook JR, Drummond M, Glick H, Heyse JF. Assessing the appropriateness of combining economic data from multinational clinical trials. *Stat Med* 2003; 22(12): 1955-1976.
61. Robins JM, Blevins D, Ritter G, Wulfsohn M. G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients. *Epidemiology* 1992; 3(4): 319-336.
62. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran ME, Berry D (Ed). *Statistical models in epidemiology: the environment and clinical trials*. New York: Springer; 1999. S. 95-134. (The IMA volumes in mathematics and its applications; Band 116).
63. Robins JM, Hernan MA, Siebert U. Estimations of the effects of multiple interventions. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL (Ed). *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*. Genf: WHO; 2004. S. 2191-2230.
64. Draper NR, Smith H. *Applied regression analysis*. New York: Wiley; 1998.
65. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999; 10(1): 37-48.
66. Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci USA* 2002; 99(Suppl 3): 7280-7287.
67. Cooper K, Brailsford SC, Davies R. Choice of modeling technique for evaluating health care interventions. *J Oper Res Soc* 2007; 58(2): 168-176.
68. Koopman JS, Jacquez G, Chick SE. New data and tools for integrating discrete and continuous population modeling strategies. *Ann N Y Acad Sci* 2001; 954: 268-294.
69. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. *Pharmacoeconomics* 2008; 26(2): 131-148.
70. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. *Pharmacoeconomics* 2006; 24(11): 1043-1053.
71. Koopman J. Modeling infection transmission. *Annu Rev Public Health* 2004; 25: 303-326.
72. Caro JJ. Pharmacoeconomic analyses using discrete event simulation. *Pharmacoeconomics* 2005; 23(4): 323-332.
73. Shechter SM, Schaefer AJ, Braithwaite RS, Roberts MS. Increasing the efficiency of Monte Carlo cohort simulations with variance reduction techniques. *Med Decis Making* 2006; 26(5): 550-553.

74. Graf von der Schulenburg JM, Greiner W, Jost F, Klusen N, Kubin M, Leidl R et al. Deutsche Empfehlungen zur gesundheitsökonomischen Evaluation: dritte und aktualisierte Fassung des Hannoveraner Konsens. *Gesundheitsökonomie & Qualitätsmanagement* 2007; 12(5): 285-290.
75. Kuntz KM, Weinstein MC. Modelling in economic evaluation. In: Drummond MF, McGuire AE (Ed). *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press; 2001. S. 141-171.
76. Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required? *Value Health* 2006; 9(4): 244-252.
77. Hollingworth W, Spackman DE. Emerging methods in economic modeling of imaging costs and outcomes a short report on discrete event simulation. *Acad Radiol* 2007; 14(4): 406-410.
78. Lee RC, Donaldson C, Cook LS. The need for evolution in healthcare decision modeling. *Med Care* 2003; 41(9): 1024-1033.
79. Banks J, Carsons JS, Nelson BL, Nicol DM. *Discrete-event system simulation*. New Jersey: Pearson Prentice Hall; 2005.
80. Macal CM, North MJ. Tutorial on agent-based modeling and simulation [online]. In: Kuhl ME, Steiger NM, Armstrong FB, Joines JA (Ed). *Proceedings of the 2005 Winter Simulation Conference*; 04.-07.12.2005; Orlando, USA. 2005. S. 2-15. [Zugriff: 09.10.2009]. URL: <http://www.informs-sim.org/wsc05papers/002.pdf>.
81. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making* 1983; 3(4): 419-458.
82. Briggs AH, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998; 13(4): 397-409.
83. Siebert U, Mühlberger N, Schöffski O. Evidenzsynthese: Meta-Analysen und Entscheidungsanalysen. In: Graf von der Schulenburg JM (Ed). *Gesundheitsökonomische Evaluationen*. Berlin: Springer; 2007. S. 261-310.
84. Hunink MG, Glasziou PP, Siegel JE, Weeks JC, Pliskin JS, Elstein AS et al. *Decision making in health and medicine: integrating evidence and values*. Cambridge: Cambridge University Press; 2001.
85. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ* 2003; 12(10): 837-848.
86. Davies HTO, Davies R. Simulating health systems: modelling problems and software solutions. *Eur J Oper Res* 1995; 87(1): 35-44.
87. Fishman GS. *Discrete-event simulation: modeling, programming, and analysis*. Berlin: Springer; 2001.

88. Bratley P, Fox BL, Schrage LE. A guide to simulation. New York: Springer; 1987.
89. Kelton WD, Sadowski RP, Sturrock DT. Simulation with ARENA. Boston: McGraw-Hill; 2007.
90. Alexopoulos C, Goldsman D, Fontanesi J, Sawyer M, De Guire M, Kopald D et al. A discrete-event simulation application for clinics serving the poor [online]. In: Peters BA, Smith JS, Medeiros DJ, Rohrer MW (Ed). Proceedings of the 2001 Winter Simulation Conference; 09.-12.09.2001; Arlington, USA. 2001. S. 1386-1391. [Zugriff: 09.10.2009]. URL: <http://www.informs-sim.org/wsc01papers/188.PDF>.
91. Cahill W, Render M. Dynamic simulation modeling of ICU bed availability [online]. In: Farrington PA, Nembhard HB, Sturrock GW, Evans GW (Ed). Proceedings of the 1999 Winter Simulation Conference; 05.-08.12.1999; Squaw Peak, USA. 1999. S. 1573-1576. [Zugriff: 09.10.2009]. URL: <http://www.informs-sim.org/wsc99papers/228.PDF>.
92. Gray J, Geva A, Zheng Z, Zupancic JA. CoolSim: using industrial modeling techniques to examine the impact of selective head cooling in a model of perinatal regionalization. Pediatrics 2008; 121(1): 28-36.
93. Jun JB, Jacobson SH, Swisher JR. Applications of discrete-event simulation in health care clinics: a survey. J Oper Res Soc 1999; 50(2): 109-123.
94. Proctor T. Simulation in hospitals. Health Manpow Manage 1996; 22(5): 40-44.
95. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making 1993; 13(4): 322-338.
96. Duffie D, Glynn P. Estimation of continuous-time Markov processes sampled at random time intervals. Econometrica 2004; 72(6): 1773-1808.
97. Stewart WJ. Introduction to the numerical solution of Markov chains. Princeton: Princeton University Press; 1994.
98. Bilge U, Saka O. Agent-based simulations in health care. In: Hasman A, Haux R, Van der Lei J, De Clercq E, Roger France FH (Ed). Ubiquity: technologies for better health in aging societies; proceedings of MIE2006. Amsterdam: IOS Press; 2006. S. 699-704. (Studies in health technology and informatics; Band 124).
99. Burke DS, Epstein JM, Cummings DAT, Parker JI, Cline KC, Singa RM et al. Individual-based computational modeling of smallpox epidemic control strategies. Acad Emerg Med 2006; 13(11): 1142-1149.
100. Gorman DM, Mezcic J, Mezcic I, Gruenewald PJ. Agent-based modeling of drinking behavior: a preliminary model and potential applications to theory and practice. Am J Public Health 2006; 96(11): 2055-2060.
101. Anderson RM, May RM. Infectious disease of humans: dynamics and control. Oxford: Oxford University Press; 1991.

102. Bojke L, Claxton K, Palmer S, Sculpher MJ. Defining and characterising structural uncertainty in decision analytic models [online]. 03.2006 [Zugriff: 09.10.2009]. (CHE Research Papers; Band 9). URL: <http://www.york.ac.uk/inst/che/pdf/rp9.pdf>.
103. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 2000; 17(5): 479-500.
104. Cantor SB, Ganiats TG. Incremental cost-effectiveness analysis: the optimal strategy depends on the strategy set. *J Clin Epidemiol* 1999; 52(6): 517-522.
105. Karlsson G, Johannesson M. Cost-effectiveness analysis and capital costs. *Soc Sci Med* 1998; 46(9): 1183-1191.
106. O'Hagan A, McCabe C, Akehurst R, Brennan A, Briggs A, Claxton K et al. Incorporation of uncertainty in health economic modelling studies. *Pharmacoeconomics* 2005; 23(6): 529-536.
107. Law AM, Kelton WD. *Simulation modelling and analysis*. New York: McGraw Hill; 2000.
108. Ross SM. *Simulation*. Oxford: Elsevier Academic Press; 2006.
109. Drummond MF, Brandt A, Luce B, Rovira J. Standardizing methodologies for economic evaluation in health care: practice, problems, and potential. *Int J Technol Assess Health Care* 1993; 9(1): 26-36.
110. Collège des Économistes de la Santé. French guidelines for the economic evaluation of health care technologies [online]. 09.2004 [Zugriff: 09.10.2009]. URL: http://www.ces-asso.org/docs/France_Guidelines_HE_Evaluation.PDF.
111. Oberkampf WL, DeLand SM, Rutherford BM, Diegetd KV, Alvine KF. Error and uncertainty in modeling and simulation. *Reliability Engineering and System Safety* 2002; 75(3): 333-357.
112. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves-facts, fallacies and frequently asked questions. *Health Econ* 2004; 13(5): 405-415.
113. Law AM, McComas MG. How to build valid and credible simulation models [online]. In: Peters BA, Smith JS, Medeiros DJ, Rohrer MW (Ed). *Proceedings of the 2001 Winter Simulation Conference*; 09.-12.09.2001; Arlington, USA. 2001. S. 22-29. [Zugriff: 09.10.2009]. URL: <http://www.informs-sim.org/wsc01papers/004.PDF>.
114. Sargent RG. Validation and verification of simulation models [online]. In: Ingalls RG, Rossetti MD, Smith JS, Peters BA (Ed). *Proceedings of the 2004 Winter Simulation Conference*; 05.-08.12.2004; Washington, USA. 2004. S. 17-28. [Zugriff: 09.10.2009]. URL: <http://www.informs-sim.org/wsc04papers/004.pdf>.