

Entscheidungsanalytische Modelle zur Sicherung der Übertragbarkeit internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems

Ein Methodenbeitrag zu HTA

Uwe Siebert

Health Technology Assessment

Herausgeber:

Deutsche Agentur für Health Technology Assessment des
Deutschen Instituts für Medizinische Dokumentation und Information
(DAHTA@DIMDI)



In der elektronischen Zeitschrift gms Health Technology Assessment der Deutschen Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA@DIMDI) werden Forschungsbeiträge, Untersuchungen, Umfragen usw. als Diskussionsbeiträge im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung veröffentlicht. Die Verantwortung für den Inhalt obliegt den jeweiligen Autoren bzw. der jeweiligen Autorin / Autor.

Deutsche Agentur für Health Technology Assessment des
Deutschen Instituts für Medizinische Dokumentation und Information
(DAHTA@DIMDI)

Informationssystem
Health Technology Assessment (HTA)
in der Bundesrepublik Deutschland

**Entscheidungsanalytische Modelle zur Sicherung der
Übertragbarkeit internationaler Evidenz von HTA auf
den Kontext des deutschen Gesundheitssystems**

Ein Methodenbeitrag zu HTA

U. Siebert

© DAHTA@DIMDI. Alle Rechte, auch die des Nachdrucks von Auszügen, der photomechanischen Wiedergabe und der Übersetzung, vorbehalten.

Schriftenreihe Health Technology Assessment, Bd. 16

ISSN: 1864-9645

1. Auflage

DIMDI

Waisenhausgasse 36 - 38a

50676 Köln

Telefon: 0221 / 4724 - 1

Telefax: 0221 / 4724 - 444

Inhaltsverzeichnis

1	Zusammenfassung	1
2	Wissenschaftliche Zusammenfassung	2
3	Hauptdokument	5
3.1	Gesundheitspolitische Fragestellung	5
3.2	Einführung	5
3.3	Forschungsfragen	7
3.4	Methodik.....	8
3.4.1	Systematische Literatursuche zu Problemen und Lösungsansätzen im Rahmen entscheidungsanalytischer Evidenzübertragung	8
3.4.2	Methoden der Entscheidungsanalyse.....	9
3.5	Ergebnisse	46
3.5.1	Systematische Literaturrecherche	46
3.5.2	Literaturbeispiele zu Standardsituationen notwendiger Evidenzübertragung .	49
3.5.3	Zentrale Versorgungsbereiche und Parameter bei der Evidenzübertragung ..	53
3.5.4	Entscheidungsanalyse als Instrument zur Evidenzübertragung.....	62
3.5.5	Stufenkonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext.....	64
3.6	Diskussion	68
3.6.1	Literaturrecherche.....	68
3.6.2	Parameterliste und Stufenkonzept.....	69
3.6.3	Übertragung internationaler Evidenz auf den nationalen Kontext	69
3.6.4	Standardsituationen für den Einsatz von Entscheidungsanalysen.....	72
3.6.5	Limitationen und kritische Würdigung des entscheidungsanalytischen Ansatzes	75
3.7	Schlussfolgerungen und Empfehlungen	78
4	Anhang	80
4.1	Literatur	80
4.2	Abbildungsverzeichnis	87
4.3	Tabellenverzeichnis	88
4.4	Abkürzungsverzeichnis	89

Vorbemerkungen

Die Deutsche Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA@DIMDI) hat Herrn Uwe Siebert, MPH, MSc, mit der Erstellung dieses Berichts beauftragt.

Das interne Gutachten, eine inhaltliche und eine formale Kontrolle der Arbeit, wurde von Mitarbeitern von DAHTA@DIMDI durchgeführt.

Die externe Begutachtung erfolgte durch Prof. Dr. C. Weihs, Lehrstuhl Computergestützte Statistik, Fachbereich Statistik, Universität Dortmund. Die Gutachter tragen keine Verantwortung für den Inhalt des Berichts.

Die Erstellung, Begutachtung und Überarbeitung des Berichts erfolgte auf der Grundlage standardisierter Arbeitsanweisungen. Für alle an der Erstellung beteiligten Personen bestanden keine finanziellen oder sonstigen Interessenkonflikte.

Die Basis der Finanzierung des Gesamtberichts bildet der gesetzliche Auftrag nach Artikel 19 des GKV-Gesundheitsreformgesetzes 2000 und erfolgte durch die Deutsche Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA@DIMDI) im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung.

1 Zusammenfassung

(Verfasser: U. Siebert)

Ziele dieses HTA-Methodenberichts (HTA = Health Technology Assessment) sind die Prüfung und die Bewertung der Entscheidungsanalyse (EA) als methodisches Instrument zur Übertragung sowie zur Anpassung internationaler Evidenz aus dem Bereich HTA auf den Kontext des deutschen Gesundheitssystems. Ferner soll ein systematisches Konzept entwickelt werden, das als Orientierung bei der Auswahl, bei der Übertragung, bei der Anpassung und bei der Zusammenführung dieser Daten im Rahmen deutscher HTA-Projekte dienen kann.

Es soll ein Überblick über die Konzepte und Methoden der EA gegeben sowie neben den Anwendungsbereichen, den Möglichkeiten und den Grenzen entscheidungsanalytischer Verfahren die Grundlage für eine kritische Beurteilung entscheidungsanalytischer Studien vermittelt werden. Die beiden wichtigsten Methoden der EA sind das Entscheidungsbaumverfahren und Markov-Modelle. Zusätzlich werden die verschiedenen Formen von Sensitivitätsanalysen dargestellt. Anhand von Beispielen werden Standardsituationen für die Anwendung von EA zur Evidenzübertragung erläutert und zentrale Versorgungsbereiche und Parameter bei der Evidenzübertragung aufgezeigt. Schließlich soll ein Stufenkonzept (Framework) zur Übertragung internationaler Evidenz im Rahmen von HTA auf den nationalen Kontext entwickelt sowie die Stärken und die Limitationen des entscheidungsanalytischen Ansatzes kritisch gewürdigt und herausgearbeitet werden.

Zusammenfassend können in diesem HTA-Bericht verschiedene Situationen, in denen sich der Einsatz entscheidungsanalytischer Modelle anbietet, dargestellt und der Nutzen des Einsatzes von EA im Rahmen der Übertragung internationaler Evidenz auf den nationalen Kontext aufgezeigt werden. Es wird ein systematisches Instrumentarium zur Übertragung internationaler Evidenz auf den nationalen Versorgungskontext entwickelt, das in zwei HTA-Projekten erfolgreich angewandt und in weiteren HTA-Projekten, bei denen sich die Problematik der Übertragung von Evidenz aus ausländischen Studien auf den deutschen Kontext stellt, empfohlen wird. Der Einsatz entscheidungsanalytischer Modelle zur Evidenzübertragung wird befürwortet. Er besitzt jedoch verschiedene Limitationen, die in jedem HTA-Bericht transparent dargestellt und kritisch diskutiert werden sollen.

2 Wissenschaftliche Zusammenfassung

(Verfasser: U. Siebert)

Ziel

Ziele dieses HTA-Methodenberichts sind die Prüfung und die Bewertung der Entscheidungsanalyse (EA) als methodisches Instrument zur Übertragung und zur Anpassung internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems. Ferner soll ein systematisches Konzept entwickelt werden, das als Orientierung bei der Auswahl, bei der Übertragung, bei der Anpassung und bei der Zusammenführung dieser Daten im Rahmen deutscher HTA-Projekte dienen kann.

Hintergrund

Durch den immer schneller voranschreitenden medizinischen Fortschritt und die dadurch bedingte Kostensteigerung im Gesundheitssystem kommt den Verfahren zur Bewertung der klinischen und ökonomischen Konsequenzen von Medizintechnologien eine zentrale Bedeutung zu. Für den Nachweis der klinischen Wirksamkeit (Efficacy) einer Technologie gilt die kontrollierte klinische Studie als Goldstandard. Studien dieses Typs sind jedoch ausgesprochen zeitaufwendig und kostenintensiv sowie für viele Technologien aus praktischen oder ethischen Gründen nicht durchführbar. Darüber hinaus gewinnt unter der Knappheit finanzieller Ressourcen im Gesundheitssystem die Beurteilung der ökonomischen Konsequenzen neuer Verfahren immer mehr an Bedeutung. Für die nationalen Entscheidungsträger des Gesundheitssystems ergibt sich häufig das Problem, dass zur Beantwortung drängender Fragestellungen notwendige Daten zwar aus internationalen Studien, aber nicht aus dem nationalen Versorgungskontext vorliegen, die Übertragung internationaler Ergebnisdaten fraglich und die valide Erhebung nationaler Daten neben der Verursachung hoher Kosten vor allem sehr lange Zeit beansprucht.

Der Einsatz entscheidungsanalytischer Modelle ermöglicht unter bestimmten Annahmen die flexible Verknüpfung verschiedener evidenzbasierter Datenbestände. Sowohl die Übertragung der evidenzbasierten Daten aus internationalen Studien auf den deutschen Kontext, als auch die Synthese klinischer und ökonomischer Daten zur Beurteilung der Wirtschaftlichkeit einer Technologie im nationalen Kontext sowie schließlich die Zusammenführung von Ergebnissen aus klinischen Studien mit intermediären Endpunkten und epidemiologischen Beobachtungsstudien mit Langzeitendpunkten erfordert die Verwendung entscheidungsanalytischer Modelle.

Fragestellungen

Folgende Fragestellungen werden im Rahmen dieses Berichts bearbeitet:

1. Welches sind Standardsituationen, bei denen internationale Evidenz auf den Kontext eines bestimmten Gesundheitssystems zu übertragen ist?
2. Welche Parameter sind für entscheidungsanalytische Modellierungen zu berücksichtigen?
3. Welche strategischen Vorgehensweisen (Framework) erscheinen im Rahmen deutscher HTA-Projekte sinnvoll?
4. Welche Möglichkeiten bietet die EA bei der Synthese und Übertragung von Evidenz?
5. Welchen Limitationen unterliegt die EA als Instrument zur Übertragung von Evidenz?

Methoden

Zur Identifikation relevanter Literatur zu Problemen und zu Lösungsansätzen im Rahmen entscheidungsanalytischer Evidenzübertragung wird eine systematische Literaturrecherche durchgeführt. Die Anwendungsbeispiele sollen einen Vergleich zwischen verschiedenen Ländern oder Regionen im Hinblick auf klinische, epidemiologische oder gesundheitsökonomische Fragestellungen aufweisen.

Es soll ein Überblick über die Konzepte und Methoden der EA gegeben und neben den Anwendungsbereichen, Möglichkeiten und Grenzen entscheidungsanalytischer Verfahren die Grundlage für eine kritische Beurteilung entscheidungsanalytischer Studien vermittelt werden. Die wichtigsten Methoden der EA (Entscheidungsbaumverfahren und Markov-Modelle) sowie die verschiedenen Formen der Sensitivitätsanalysen werden dargestellt.

Anhand von Beispielen werden Standardsituationen für die Anwendung von EA zur Evidenzübertragung erläutert sowie zentrale Versorgungsbereiche und Parameter bei ihr aufgezeigt.

Schließlich sollen ein Stufenkonzept (Framework) zur Übertragung internationaler Evidenz im Rahmen von HTA auf den nationalen Kontext entwickelt sowie die Stärken und die Limitationen des entscheidungsanalytischen Ansatzes kritisch gewürdigt und herausgearbeitet werden.

Ergebnisse und Schlussfolgerungen

Die Ergebnisse der Literaturrecherche zeigen eine Diskrepanz zwischen den Lehrbuchempfehlungen, die den entscheidungsanalytischen Ansatz bei der Übertragung von Evidenz zwischen verschiedenen Gesundheitssystemen propagieren und der Praxis, die durch einen Mangel an Anwendungsbeispielen in der Literatur gekennzeichnet ist. Diese Lücke verdeutlicht den Forschungsbedarf zu diesem Thema.

Zusammenfassend können in diesem HTA-Bericht verschiedene Situationen, in denen sich der Einsatz entscheidungsanalytischer Modelle anbietet, dargestellt und der Nutzen des Einsatzes von EA im Rahmen der Übertragung internationaler Evidenz auf den nationalen Kontext aufgezeigt werden.

Es wird ein systematisches Instrumentarium zur Übertragung internationaler Evidenz auf den nationalen Versorgungskontext vorgestellt. Dieses Instrumentarium beinhaltet eine Liste der bei der Evidenzübertragung zu prüfenden Parameter, die aus den folgenden Bereichen stammen: Soziodemographie, Epidemiologie, medizinische Outcomes, Patientenpräferenzen / Lebensqualität, Gesundheitsökonomie, Versorgungs- und Technologiestandards, Struktur des Gesundheitssystems sowie weitere Determinanten.

Ferner wird ein Stufenkonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext entwickelt, das im Einzelnen konkrete und praktische Schritte beschreibt, wie bei der Durchführung der entscheidungsanalytischen Evidenzübertragung vorzugehen ist und welche Optionen es dabei für verschiedene gegebene Rahmenbedingungen gibt. Das Stufenkonzept besteht aus den folgenden Punkten:

1. Identifikation internationaler Evidenz.
2. Identifikation der am Entscheidungsproblem beteiligten und relevanten Parameter.
3. Bewertung der Übertragbarkeit der internationalen Ergebnisse anhand einer Prüfliste.
4. Entscheidung, ob eine Reanalyse für den deutschen Kontext notwendig ist.
5. Entscheidung, ob eine EA sinnvoll / notwendig ist.

6. Aufbau eines interdisziplinären Expertennetzwerks.
7. Prüfung der Datenverfügbarkeit für Deutschland und die sich daraus ergebende Datenerhebung.
8. Modellentwicklung oder strukturelle Modelladaptation.
9. Numerische Implementation oder Adaptation der Modellparameter.
10. Klinische und / oder Kosteneffektivitätsanalyse.
11. Sensitivitätsanalysen.
12. Ergebnisinterpretation und Schlussfolgerungen / Empfehlungen.
13. Festlegung des weiteren Forschungsbedarfs.

Dieses Stufenkonzept wird bereits im Rahmen von zwei deutschen HTA-Berichten in den Bereichen der antiviralen Therapie der chronischen Hepatitis C und der Zervixkarzinomfrüherkennung eingesetzt. Es erweist sich als praktikabel und als erfolgreich. Der Einsatz des entwickelten Stufenkonzepts in weiteren HTA-Projekten, bei denen sich die Problematik der Übertragung von Evidenz aus ausländischen Studien auf den deutschen Kontext stellt, wird empfohlen. Diese Verwendung entscheidungsanalytischer Modelle besitzt jedoch verschiedene Limitationen, die in jedem HTA-Bericht transparent dargestellt und kritisch diskutiert werden sollten. Eine Anwendung von EA in zukünftigen HTA-Projekten muss zeigen, ob sich der Ansatz grundsätzlich für die Evidenzübertragung eignet sowie welche Modifikationen und Erweiterungen dafür möglicherweise erforderlich sind.

3 Hauptdokument

3.1 Gesundheitspolitische Fragestellung

(Verfasser: U. Siebert)

Im Rahmen von Health Technology Assessment (HTA) ist es in vielen Fällen nicht möglich, die Bewertung der zu untersuchenden Technologie ausschließlich auf Daten des Zielkontexts, sowie der Zielpopulation, der Region, oder des Gesundheitssystems zu basieren.

Entscheidungsanalytische Modelle dienen der formalen Zusammenführung der Evidenz aus verschiedenen Bereichen, so dass sie eingesetzt werden können, um die Forschungsfragen im Rahmen eines HTA-Berichts in systematischer, expliziter und transparenter Weise zu beantworten. Sowohl die Übertragung der evidenzbasierten Daten aus internationalen Studien auf den deutschen Kontext, als auch die Synthese klinischer und ökonomischer Daten zur Beurteilung der Wirtschaftlichkeit einer Technologie im nationalen Kontext sowie schließlich die Zusammenführung von Ergebnissen aus klinischen Studien mit intermediären Endpunkten und epidemiologischen Beobachtungsstudien mit Langzeitendpunkten erfordert den Einsatz entscheidungsanalytischer Modelle.

Für deutsche HTA-Berichte wird gefordert, dass neben systematischen Übersichtsarbeiten zusätzlich gesundheitsökonomische Berechnungen und mathematische Modellanalysen durchgeführt werden²⁷, um dem Anspruch eines HTA-Berichts im Sinn einer systematischen Evaluation von medizinischer Effektivität und Wirtschaftlichkeit medizinischer Technologien gerecht zu werden.

Ziele dieses HTA-Methodenprojekts sind die Prüfung und die Bewertung der EA als methodisches Instrument zur Übertragung und Anpassung internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems. Ferner soll ein systematisches Konzept entwickelt werden, das als Orientierung bei der Auswahl, bei der Übertragung, bei der Anpassung und bei der Zusammenführung dieser Daten im Rahmen deutscher HTA-Projekte dienen kann.

3.2 Einführung

(Verfasser: U. Siebert)

Durch den immer schneller voranschreitenden medizinischen Fortschritt und die dadurch bedingte Kostensteigerung im Gesundheitssystem kommt den Verfahren zur Bewertung der klinischen und der ökonomischen Konsequenzen von Medizintechnologien eine zentrale Bedeutung zu. Durch die evidenzbasierte Gesundheitsversorgung²⁷ werden neue Wege des zeitnahen Wissens- und Informationsmanagements ermöglicht, das mit der Etablierung von HTA auch in Deutschland seinen Eingang als Instrument zur Bereitstellung von systematisch aufbereiteter Information für die Entscheidungsträger im Gesundheitssystem gefunden hat²⁷.

HTA befasst sich mit der umfassenden Synthese sowie mit der Bewertung aller relevanten Kenntnisse und Aspekte bezüglich einer medizinischen Technologie, die für Entscheidungen im Gesundheitswesen und auf der Ebene der Bedarfsplanung notwendig sind. Hierzu gehören vor allem sowohl die systematische Untersuchung kurz- und langfristiger klinischer Konsequenzen der Anwendung einer medizinischen Technologie, als auch die Evaluation der ökonomischen Kurz- und Langzeitkonsequenzen beim Einsatz der untersuchten Technologie im Kontext des deutschen Gesundheitssystems.

Für den Nachweis der klinischen Wirksamkeit (Efficacy) einer Technologie gilt die randomisierte

kontrollierte klinische Studie (RCT) als Goldstandard². Studien dieses Typs sind jedoch ausgesprochen zeitaufwendig und kostenintensiv. Weitere Limitationen liegen in der eingeschränkten Aussagekraft von RCT bezüglich der Effektivität (Effectiveness) der Technologie unter Routinebedingungen und der oft unklaren Übertragbarkeit auf andere als den in der RCT untersuchten Studienpopulationen (z.B. unterschiedliches Alter, Komorbidität, Krankheitschweregrad). Diese Nachteile gelten im besonderen Maß für international angelegte Multicenterstudien. Schließlich wird im RCT meist nur eine einzige Wirksamkeitshypothese untersucht, die sich in vielen Fällen auf einen intermediären Kurzzeitendpunkt bezieht.

Darüber hinaus gewinnt unter der Knappheit finanzieller Ressourcen im Gesundheitssystem die Beurteilung der ökonomischen Konsequenzen neuer Verfahren immer mehr an Bedeutung. Im Zuge dieser Entwicklung werden im Rahmen von RCT mit zunehmender Häufigkeit gesundheitsökonomische Daten mit erfasst („Piggy Back“ Studien)⁶¹. Aber auch die Ergebnisse dieser Arbeiten unterliegen den Limitationen der Übertragbarkeit auf den Routinekontext des nationalen Gesundheitssystems und der fehlenden Aussagekraft bezüglich eines langfristigen Zeithorizonts.

Für die nationalen Entscheidungsträger des Gesundheitssystems ergibt sich häufig das Dilemma, dass zur Beantwortung drängender Fragestellungen notwendige Daten zwar aus internationalen Studien, aber nicht aus dem nationalen Versorgungskontext vorliegen, und die valide Erhebung dieser Daten neben der Verursachung hoher Kosten vor allem sehr lange Zeit beanspruchen würde. Gerade vor dem Hintergrund einer sich zumindest in Teilbereichen rasch verändernden Landschaft von Diagnoseverfahren und Therapiemaßnahmen, wie beispielsweise bei der Behandlung von AIDS-Kranken⁵⁶ erscheint es nahezu unmöglich, evidenzgestützte Entscheidungen zeitnah zu treffen. Eine gegenläufige Entwicklung, die jedoch auf die gleiche Fragestellung hinführt, besteht in der exponentiell anwachsenden Flut von Empfehlungen, Leit- und Richtlinien. So gibt es für britische Allgemeinärzte 1999 etwa 850 verschiedene Leitlinien⁶⁴.

Bei den medizinischen, politischen und planerischen Entscheidungen im Gesundheitssystem sind also verschiedene Aspekte zu berücksichtigen: (i) sowohl klinische, ökonomische und planerische Effekte, (ii) neben den intermediären Endpunkten zusätzlich finale Endpunkte wie Mortalität und Lebensqualität, (iii) die Aktualität der Datenlage sowie (iv) die Besonderheiten des nationalen Versorgungssystems, auf das die Evidenz übertragen werden muss.

Aus den genannten Gründen sollen HTA-Berichte neben systematischen Übersichtsarbeiten zusätzlich gesundheitsökonomische Berechnungen und mathematische Modellanalysen enthalten²⁷. Letztere können sich auf klinische, ökonomische oder planerische Aspekte beziehen. Beispielsweise interessiert bei der klinischen Evaluation der antiviralen Therapie der Hepatitis C nicht nur der in internationalen „Multicenter-RCT“ evaluierte Effekt auf die Virusfreiheit nach Therapie, sondern auch der langfristige Therapieeffekt auf die Lebenserwartung (Mortalität) und -qualität der Patienten. Analog hat eine gesundheitsökonomische Evaluation nicht nur die Arzneimittelkosten der maximal einjährigen antiviralen Therapie zu berücksichtigen, sondern ebenso die therapiebedingten Einsparungen durch eine Reduktion der Hepatitis C-bedingten Folgeerkrankungen und -ereignisse wie Komplikationen durch dekompensierte Leberzirrhosen, Leberzellkarzinome oder Lebertransplantationen. Schließlich sollen die bedarfsplanerischen Aspekte systematisch und quantitativ untersucht werden, wie beispielsweise ein Vergleich der innerhalb der nächsten zehn Jahre insgesamt in Deutschland erforderlichen Lebertransplantationen bei den Szenarien mit und ohne antiviraler Therapie. Diese können dann der Anzahl verfügbarer Organtransplantate gegenübergestellt werden.

Damit besteht von verschiedenen Seiten ein starkes Interesse an Methoden zur kostengünstigen und zur zeitnahen Übertragung bzw. Anpassung evidenzbasierter und interdisziplinärer Daten, die für andere Länder bzw. Regionen vorliegen oder sogar für andere Fragestellungen erhoben werden, und die im Kontext des konkret entscheidungsbedürftigen Sachverhalts zusammengeführt werden sollen.

Das Verfahren der EA, das sich in den letzten Jahren einer zunehmenden Verbreitung erfreut, bietet sich für eine derartige Fragestellung an⁶¹. Der Einsatz entscheidungsanalytischer Modelle ermöglicht unter bestimmten Annahmen die flexible Verknüpfung verschiedener evidenzbasierter Datenbestände. Sowohl die Übertragung der evidenzbasierten Daten aus internationalen Studien auf den deutschen Kontext, als auch die Synthese klinischer und ökonomischer Daten zur Beurteilung der Wirtschaftlichkeit einer Technologie im nationalen Kontext sowie schließlich die Zusammenführung von RCT-Ergebnissen mit intermediären Endpunkten und epidemiologischen Beobachtungsstudien mit Langzeitendpunkten erfordert den Einsatz entscheidungsanalytischer Modelle. Einmal modellierte Sachverhalte können, sofern sich nicht grundlegende strukturelle Änderungen ergeben, bei neuer Evidenz mit geringerem Zeit- und Kostenaufwand aktualisiert werden.

Obwohl in anderen Ländern die Durchführung von systematischen EA für viele HTA-Institutionen bereits Standard ist und auch für deutsche HTA-Berichte in den von der German Scientific Working Group Technology Assessment for Health Care entwickelten Standards für deutsche HTA-Berichte im Auftrag des Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI) vorgesehen sind, werden EA bislang nur vereinzelt in HTA-Berichten eingesetzt^{26, 18, 67, 68, 80, 81, 82}.

Ziele dieses HTA-Methodenprojekts sind die Prüfung und die Bewertung der EA als methodisches Instrument zur Übertragung und Anpassung internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheitssystems. Ferner soll ein systematisches Konzept entwickelt werden, das als Orientierung bei der Auswahl, bei der Übertragung, bei der Anpassung und bei der Zusammenführung dieser Daten im Rahmen deutscher HTA-Projekte dienen kann.

3.3 Forschungsfragen

(Verfasser: U. Siebert)

Im Rahmen der Durchführung von HTA-Projekten liegt teils hochwertige Evidenz aus anderen Ländern oder einem verwandten Problemkontext vor. Dahingegen sind nationale Daten entweder unvollständig oder fehlen vollständig bzw. sind von eingeschränkter oder mangelhafter Qualität. Ferner gibt es Situationen, in denen bereits entscheidungsanalytische Modellierungen zur Beantwortung gesundheitsbezogener Fragen und der Bewertung der Implementierung medizinischer Technologien für andere Länder durchgeführt wurden.

In den seltensten Fällen können die vorliegenden Primärdaten oder entsprechende Modellierungsergebnisse aus anderen Ländern direkt und ohne Anpassungen übernommen werden. Exemplarisch sei auf die unterschiedlichen Häufigkeitsverteilungen von Erkrankungen in verschiedenen Ländern oder sogar regional verschieden verteilte Risikofaktoren hingewiesen. Auch das medizinische Vorgehen (Versorgungspraxis) und Kostenparameter unterscheiden sich meist zwischen verschiedenen Ländern und ihren Gesundheitssystemen. Aus diesem Grund soll der Einsatz entscheidungsanalytischer Methoden im Rahmen der Übertragung und Anpassung internationaler Evidenz von HTA auf den Kontext des deutschen Gesundheits-

systems geprüft und bewertet werden.

Daraus ergeben sich folgende Teilfragestellungen:

1. Welches sind Standardsituationen, bei denen internationale Evidenz auf den Kontext eines bestimmten Gesundheitssystems zu übertragen ist?
2. Welche Parameter sind für entscheidungsanalytische Modellierungen zu berücksichtigen?
3. Welche strategischen Vorgehensweisen (Framework) erscheinen im Rahmen deutscher HTA-Projekte sinnvoll?
4. Welche Möglichkeiten bietet die EA bei der Synthese und Übertragung von Evidenz?
5. Welchen Limitationen unterliegt die EA als Instrument zur Übertragung von Evidenz?

Unterstützend wird eingangs recherchiert, ob es bereits Beispiele in der Literatur gibt, bei denen entscheidungsanalytische Modellierungen zur Übertragung medizinischer oder ökonomischer Evidenz auf andere Gesundheitssysteme eingesetzt werden.

Da es sich bei der vorliegenden Arbeit um einen methodischen HTA-Beitrag handelt, ist das Ziel der Literaturrecherchen zu den einzelnen Fragestellungen nicht die vollständige und erschöpfende Erfassung der Evidenz zu dem jeweiligen Fragenkomplex, sondern Ziel ist die Darlegung eines Methodenkonzepts. Die Literaturrecherche dient dem Auffinden von Hintergrundinformationen und dazu, die methodischen Konzepte anhand ausgewählter Beispiele aus der Literatur zu illustrieren.

3.4 Methodik

(Verfasser: U. Siebert, G. Sroczynski)

3.4.1 Systematische Literatursuche zu Problemen und Lösungsansätzen im Rahmen entscheidungsanalytischer Evidenzübertragung

Zur Identifikation relevanter Literatur zu Problemen und zu Lösungsansätzen im Rahmen entscheidungsanalytischer Evidenzübertragung wird eine systematische Literaturrecherche durchgeführt. Die Anwendungsbeispiele sollen einen Vergleich von klinischen, epidemiologischen oder gesundheitsökonomischen Fragestellungen zwischen verschiedenen Ländern oder Regionen aufweisen. Sie sollen Aussagen zur Vergleichbarkeit bzw. zur Übertragbarkeit der Daten enthalten sowie das entscheidungsanalytische Methodeninstrumentarium verwenden.

Die Recherchen werden in den Literaturdatenbanken MEDLINE™ und EconLit™ durchgeführt. Die Abfragen erfolgen unter der OVID™ Bedienungsfläche. Die Suche wird auf den Veröffentlichungszeitraum zwischen 1985 und Oktober 2001 (MEDLINE) bzw. Dezember 2001 (EconLit) eingeschränkt.

Ergänzend zu den oben beschriebenen Datenbankrecherchen wird für den spezifisch deutschen Kontext (gleicher Zeitraum wie oben) eine weitere Abfrage zur entscheidungsanalytischen Methodik in der Datenbank MEDLINE™ durchgeführt.

Die Ergebnisse werden im nächsten Schritt einzeln durchgesehen und auf ihre Relevanz überprüft.

Die aus den Datenbankrecherchen resultierende Literaturliste wird durch Hinweise aus den Literaturverzeichnissen der einschlägigen Artikel um einzelne Arbeiten ergänzt. Zur Vervollständigung erfolgt eine Durchsicht ausgewählter gesundheitsökonomischer Standardlehrbücher

und gesundheitsökonomischer Lehrbuchkapitel^{10, 24, 28, 61, 62}.

Zusätzlich werden typische Situationen exemplarisch anhand von Literaturbeispielen beschrieben, die für Entscheidungsträger im Hinblick auf eine Evidenzübertragung auf deutsche Verhältnisse von besonderem Interesse sein dürften. Die Beispiele hierzu basieren nicht auf einer umfassenden und systematischen Literaturrecherche, sondern ergeben sich aus eigenen Literatursammlungen und den im Rahmen der systematischen Literaturrecherche aufgefundenen Querverweisen.

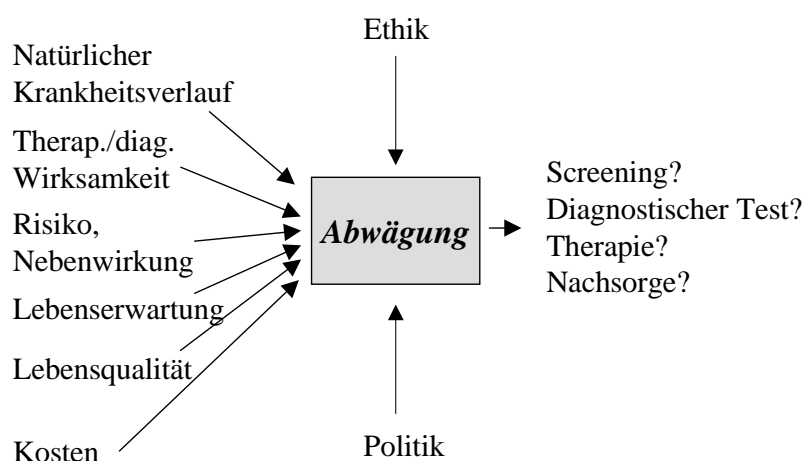
3.4.2 Methoden der Entscheidungsanalyse

3.4.2.1 Grundlagen

Steht man vor der Entscheidung für oder gegen ein bestimmtes Public Health-Programm oder soll eines unter vielen möglichen konkurrierenden Public Health-Programmen zur Prävention oder der Behandlung einer Krankheit ausgewählt werden, spielen verschiedene Kriterien wie die Verlängerung der Lebenserwartung, die Senkung der Morbidität und die Erhöhung der Lebensqualität eine Rolle. Im gesamtgesellschaftlichen Kontext sind ferner die Wirtschaftlichkeit der Maßnahmen, die Verteilungsgerechtigkeit innerhalb der Gesellschaft und der gesundheitspolitischen Ziele zu berücksichtigen (Abbildung 1). Nicht immer sind die zu erwartenden Konsequenzen einer Public Health-Maßnahme so gelagert, dass die Auswahl der optimalen Strategie leicht fällt. Häufig ist abzuwägen zwischen dem medizinischen Nutzen und den Risiken einer medizinischen Maßnahme oder einem Gewinn an Lebenserwartung und einem Verlust an Lebensqualität (bzw. umgekehrt). Geht man von einer Knappheit der Mittel im Gesundheitswesen aus³⁸, stellt sich in anderen Fällen die Frage, ob der mit einer medizinischen Maßnahme verbundene Gewinn an Gesundheit die dafür aufzuwendenden Ressourcen rechtfertigt oder ob mit diesen Ressourcen an anderer Stelle „mehr Gesundheit“ erzielt werden kann.

Die EA (Decision Analysis) versucht, solche Fragen in transparenter und systematischer Weise zu beantworten⁸³. Als quantitative Methode kommt ihr im Rahmen von Public Health-Fragestellungen eine bedeutsame Brückenfunktion bei der Synthese von Informationen aus Epidemiologie, Outcomes Research und Gesundheitsökonomie zu.

Abbildung 1: Einflussfaktoren beim Entscheidungsprozess.



Entscheidungen in der Medizin sind meistens unter Unsicherheit zu treffen. Dies bedeutet, dass für einen Patienten mit einem bestimmten Symptomenkomplex letztlich immer eine Ent-

scheidung bezüglich des weiteren diagnostischen oder therapeutischen Vorgehens zu treffen ist, auch wenn zum Zeitpunkt dieser Entscheidung viele Aspekte bezüglich seiner Erkrankung und der Auswirkungen bestimmter Therapien auf seinen Gesundheitszustand nicht mit vollständiger Sicherheit vorhergesagt werden können. Viele diagnostische Prozeduren, die zur Aufklärung des Krankheitsgeschehens eingesetzt werden, besitzen selbst ein medizinisches Risiko. So reicht das Spektrum der Komplikationen bei Biopsien der Prostata zur Diagnose des Prostatakarzinoms von Infektionen bis hin zu Todesfällen. Aber auch der Erfolg von therapeutischen Verfahren ist meist unsicher, da zum Zeitpunkt des Behandlungsbeginns oft noch nicht feststeht, ob bzw. bis zu welchem Grad die Therapie zu einem Heilungserfolg führt und welche therapieinduzierten Komplikationen bzw. Nebenwirkungen auftreten. So ist bei einem Patienten mit Prostatakarzinom im Frühstadium in keinem Fall sicher gewährleistet, dass die Erkrankung durch eine radikale Entfernung der Prostata und des anhängenden Gewebes (radikale Prostatektomie) vollständig und beschwerdefrei geheilt wird. Ferner kann es durch die Operation selbst zur Harninkontinenz, sexuellen Funktionsstörungen oder anderen Komplikationen kommen, was bei der Therapieentscheidung berücksichtigt werden sollte. Auf der anderen Seite kann diese Operation zur Heilung und zu einem beschwerdefreien Leben führen⁶⁵. Letztlich können meist die Auswirkungen der Therapie auf die Lebensqualität der Patienten und auch die verursachten Kosten nicht exakt vorhergesagt werden.

Da in jedem Fall – auch unter Unsicherheit – eine optimale Entscheidung für den Patienten getroffen werden muss, stellt sich die Aufgabe, unter Berücksichtigung der Unsicherheit und unter Abwägung der verschiedenen Konsequenzen die für den Patienten oder eine bestimmte Patientengruppe optimale medizinische Vorgehensweise zu identifizieren. Kurz: Die Entscheidung muss getroffen werden, auch unter Unsicherheit.

3.4.2.2 Definitionen und Abgrenzungen

Der Begriff der EA ist zunächst von dem weiter gefassten Bereich der allgemeinen medizinischen Entscheidungsfindung (Medical Decision Making) abzugrenzen. Die medizinische Entscheidungsfindung bedient sich verschiedener rationaler und systematischer Ansätze zur Entscheidungsfindung. Ihr Ziel ist die Verbesserung der Gesundheit und der klinischen Versorgung von Individuen und / oder Populationen, sowie die Unterstützung der Entscheidungsfindung im gesundheitspolitischen Bereich. Dies schließt neben der EA im engeren Sinn (auch formale EA genannt) die Anwendung verschiedener quantitativer Methoden im klinischen Alltag, in der epidemiologischen Forschung und in der Evaluation von Public Health-Maßnahmen ein. Ferner beschäftigt sich die medizinische Entscheidungsfindung mit der Untersuchung menschlicher Kognition und der Psychologie klinischen Denkens und Schlussfolgerns, sowie der Verwendung von Computern und künstlicher Intelligenz zur Entscheidungsunterstützung.

Die vorliegende Arbeit widmet sich ausschließlich der EA im engeren Sinn, die ein Teilgebiet des gesamten Bereichs der medizinischen Entscheidungsfindung ist. Die EA im engeren Sinn ist ein mathematisch-formaler Ansatz und verwendet klar festgelegte Algorithmen zur Optimierung des Ergebnisses einer Entscheidung. Deshalb wird sie auch formale EA genannt. Ein weiteres Synonym ist der Begriff entscheidungsanalytische Modellierung, der die Tatsache ausdrückt, dass in der formalen EA verschiedene Parameter der Krankheit, ihrer Behandlung und ihrer Konsequenzen innerhalb eines mathematischen Modells zusammengeführt werden. Meist wird einfach kurz von EA gesprochen. Auch im Folgenden wird unter EA die formale EA verstanden.

Die EA ist definiert als ein systematischer, expliziter und quantitativer Ansatz zur

Entscheidungsfindung unter Unsicherheit. Die Struktur des Ablaufs aller möglichen Ereignisse wird in Form von Ereignisbäumen dargestellt, die die Wahrscheinlichkeiten dieser Ereignisse mit ihren verschiedenen medizinischen und / oder ökonomischen Konsequenzen beinhalten. Zur Analyse dieser Daten werden probabilistische Modelle und Verfahren herangezogen, insbesondere das Entscheidungsbaumverfahren und die Markov-Modelle⁷⁶. Das Ziel der EA ist die Auswahl einer Handlungsstrategie nach Gewichtung der medizinischen Nutzen, Risiken und gegebenenfalls Kosten der verschiedenen Handlungsalternativen. Die EA verfolgt dabei das Prinzip der Nutzenmaximierung. Die Art des zu maximierenden Nutzens ist dabei vor der Analyse festzulegen bzw. zu operationalisieren. Dies kann ein einzelner Parameter wie beispielsweise höhere Überlebenswahrscheinlichkeit, bessere Lebensqualität oder geringere Kosten sein, oder eine gewichtete Kombination bzw. ein Verhältnis verschiedener Parameter. Der Prozess der EA legt Struktur, Elemente und Parameter des Entscheidungsproblems offen und macht sie damit einer Diskussion zugänglich. Dabei ist zu betonen, dass der EA lediglich die Funktion zukommt, den Entscheidungsprozess zu unterstützen, nicht ihn zu determinieren. Die endgültige Entscheidung hängt außerdem von komplexeren Determinanten aus dem Bereich der Politik und der Ethik ab (Abbildung 2). Insbesondere die Prämisse der Verteilungsgerechtigkeit kann zu Entscheidungen führen, die einer Nutzenmaximierung im oben angesprochenen Sinn entgegenstehen.

Eine vollständige gesundheitsökonomische EA vergleicht mindestens zwei medizinisch effektive Handlungsstrategien miteinander und berücksichtigt dabei Kosten sowie Effekte. Dabei wird immer der inkrementelle Ansatz gewählt^{24, 41, 72}.

Bei der Modellierung bzw. Analyse des Entscheidungsproblems werden zwei Verfahren unterschieden: das Entscheidungsbaumverfahren und Markov-Modelle. Das Entscheidungsbaumverfahren ist anzuwenden bei Entscheidungssituationen mit meistens kurzem Zeithorizont, in dem alle handlungsstrategiebedingten Ereignisse eintreten. Demgegenüber werden Markov-Modelle vorwiegend bei Problemen mit längerem Zeithorizont eingesetzt.

Abbildung 2: Stellenwert der Entscheidungsanalyse im Entscheidungsprozess.

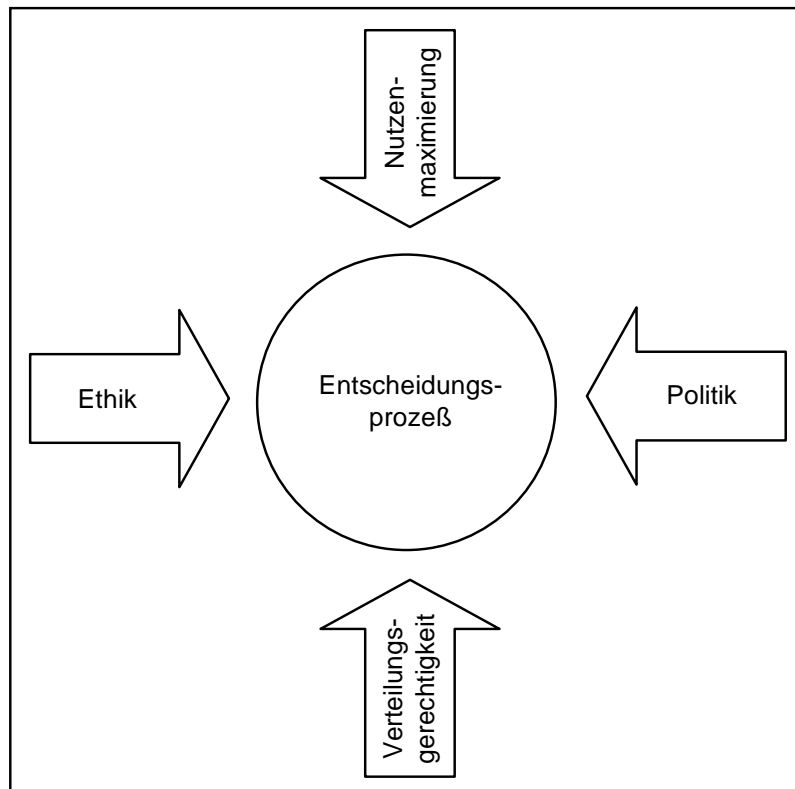


Abbildung 3 zeigt ein hypothetisches und einfaches Beispiel für einen Entscheidungsbaum. Es geht dabei um folgendes Entscheidungsproblem:

Beispiel: Diagnostischer Test

Eine sehr ernsthafte Krankheit D beginnt mit einem Symptom S. Neben der Krankheit D gibt es viele andere Ursachen für das Symptom S, die alle nicht behandlungsbedürftig sind. Auch das Symptom S ist losgelöst von der Krankheit D nicht behandlungsbedürftig. Für die Behandlung der Krankheit D steht ein Medikament zur Verfügung, das aber teils zu erheblichen Nebenwirkungen (einschließlich erhöhter Mortalität) führen kann. Es gibt einen diagnostischen Test, der bezüglich der Krankheit D mit einer gewissen Fehlerrate Kranke von Gesunden unterscheiden kann. Die Durchführung des Tests ist mit einem gewissen Letalitätsrisiko behaftet. Tritt das Symptom S bei einem Patient auf, so stehen Arzt und Patient vor der Wahl von dreierlei Handlungsalternativen:

Man beginnt sofort die medikamentöse Therapie.

Man unterlässt die medikamentöse Therapie.

Man führt den diagnostischen Test durch, der allerdings selbst ein geringes Letalitätsrisiko beinhaltet. Zudem kommt es bei diesem Test mitunter zu falschen Testergebnissen.

Dieses Muster bildet das Grundschema für jegliche (auch nichtmedizinische) EA. Viele Entscheidungsprobleme unter Unsicherheit lassen sich so strukturieren oder in Teilprobleme zerlegen, dass sich die Frage stellt, ob eine bestimmte Aktion durchgeführt, ob sie unterlassen oder ob zunächst weitere Informationen eingeholt werden sollen, um die Unsicherheit zu reduzieren. In dem (in der Praxis unrealistischen) Fall, dass das Beschaffen weiterer Information keine negativen Konsequenzen hat, ist es offensichtlich, dass diese Strategie zu wählen ist. In der Praxis ist aber oft davon auszugehen, dass das Einholen weiterer Information entweder (z.B. bei invasiven Tests) mit medizinischen Risiken oder aber zumindest mit Kosten verbunden

ist. Das Entscheidungsproblem ist damit ein Abwägen zwischen den positiven und negativen Konsequenzen des Informationsgewinns.

Bevor die einzelnen probabilistischen Verfahren und die rechnerischen Lösungsschritte für das Problem dargestellt werden, wird auf die Rahmenbedingungen einer EA eingegangen sowie der Algorithmus zur Strukturierung und Analyse eines Entscheidungsproblems unter Unsicherheit erläutert.

3.4.2.3 Design und Schritte der Entscheidungsanalyse

Bei der Durchführung einer EA sind die Rahmenbedingungen und Prämissen zu definieren, unter denen das Entscheidungsproblem analysiert werden soll. Es sind Annahmen festzulegen, unter denen das Analyseergebnis Gültigkeit haben soll. Die Festlegung der Rahmenbedingungen sowie die einzelnen Schritte, die bei der Aufstellung des entscheidungsanalytischen Modells und seiner Analyse notwendig sind, werden im Folgenden beschrieben:

1. Hintergrund und präzise Formulierung des Entscheidungsproblems.

Zunächst ist das Entscheidungsproblem genau zu spezifizieren. Es ist beispielsweise festzulegen, welche medizinischen Handlungsalternativen für welche Population (z.B. Altersgruppe) und welches Setting (z.B. ambulant oder stationär) auf Effektivität bzw. Kosteneffektivität zu prüfen und welche Effekte dabei primär zu berücksichtigen sind. Hierbei sind der medizinische und der ökonomische Problemkontext zu beachten.

2. Die Wahl der Perspektive.

Die Wahl der Perspektive hängt entscheidend von der Fragestellung ab. Beispiele für derartige Perspektiven sind:

- Patienten.
- Leistungserbringer (z.B. Kliniken, niedergelassene Ärzte).
- Kostenträger (z.B. Krankenkassen).
- Hersteller von Arzneimitteln und medizinischen Geräten.
- Arbeitgeber.
- Gesundheitsbehörden.
- Gesellschaft.

Einzelheiten zur Perspektive werden im Kapitel „Design und Schritte der Entscheidungsanalyse“ behandelt. Die Durchführung von Analysen unter verschiedenen Perspektiven ist prinzipiell möglich.

3. Festlegung des Zeithorizonts.

Es ist ein für das Problem adäquater Zeithorizont zu identifizieren, der gegebenenfalls medizinische und ökonomische Langzeitkonsequenzen abdeckt. Darunter fallen auch Rezidive oder Spätkomplikationen einer Erkrankung. Sollten sich Risiken oder Kosten im Lauf der Zeit verändern, so ist dies bei der Wahl des Zeithorizonts entsprechend zu berücksichtigen. Auch für Erkrankungen, die mehrfach auftreten können, wie beispielsweise Infektionskrankheiten oder schubartig verlaufende chronische Erkrankungen, ist ein adäquater Zeithorizont zu wählen. Ferner gibt es viele Erkrankungen, bei denen es nicht nur darum geht, ob ein medizinisches Vorgehen zur Heilung führt oder nicht, sondern die krankheitsfreie Zeit bis zum Wieder- oder Neuauftreten der Erkrankung ist wesentlich. Beispiele dafür sind Impfungen oder andere Präventivmaßnahmen und die Therapie von Krebserkrankungen. Für diese Situationen ist ein

entsprechend langer Zeithorizont zu wählen, der selbst die Konsequenzen erfasst, die zeitlich sehr weit von der eigentlichen Intervention entfernt sind.

4. Identifikation der medizinischen Handlungsalternativen.

Es sind alle relevanten Handlungsalternativen zu berücksichtigen. Dazu gehören alle Kombinationen von präventiven, diagnostischen und therapeutischen Verfahren. Oft ist eine Handlungsalternative ein festgelegter Plan bzw. Ablauf von diagnostischen und therapeutischen Maßnahmen. Dieser Plan besteht dann aus mehreren „Wenn-dann-Regeln“. Insbesondere in der schrittweisen Diagnostik werden zunächst die risikoarmen und billigeren Verfahren eingesetzt, bevor man zu den invasiveren und teureren Methoden greift. Eine a priori festgelegte Abfolge von „Wenn-dann-Regeln“ wird als Strategie bezeichnet. Insbesondere wenn mehrere diagnostische Verfahren parallel oder sequentiell eingesetzt werden können, um eine Erkrankung zu identifizieren, sind alle sinnvollen Kombinationen und Booleschen Verknüpfungen der diagnostischen Einzeltests in die Evaluation einzubeziehen.

5. Spezifizierung der möglichen klinischen Konsequenzen.

Es sind alle möglichen Gesundheitszustände festzulegen. Als Gesundheitszustand wird jeder Zustand definiert, der sich von anderen Zuständen bezüglich zukünftiger Mortalitäts-, Morbiditäts-, Risikoraten, Lebensqualität und Kosten unterscheidet, je nachdem welche Parameter in der EA berücksichtigt werden sollen. Gesundheitszustände sollen sich gegenseitig ausschließen und bezüglich der Fragestellung erschöpfend sein.

6. Darstellung des Ereignisablaufs.

Anschließend wird der Ereignisablauf schematisch dargestellt. Hierfür eignet sich beispielsweise ein Entscheidungsbaumdiagramm wie in Abbildung 3. Dieses Diagramm soll für jede der zu evaluierenden Handlungsalternativen alle möglichen Verläufe in Form von Verzweigungen abbilden. Ein durchgehender Pfad vom linken Ende (Stamm) bis zum rechten Ende (Blatt) eines Entscheidungsbaums wird als Ereignispfad bezeichnet. Am Ende des Entscheidungsbaums mündet jeder Ereignispfad in einen bestimmten Gesundheitszustand.

7. Bestimmung der Ereigniswahrscheinlichkeiten.

An jeder Verzweigung eines Entscheidungsbaums treten die nachfolgenden Ereignisse mit einer bestimmten Wahrscheinlichkeit ein, die in das Entscheidungsbaumdiagramm eingetragen wird. Einige immer wiederkehrende Beispiele für Eintrittswahrscheinlichkeiten sind:

- *Prävalenz*: Wahrscheinlichkeit für das Vorliegen eines Ereignisses.
- *Inzidenz*: Wahrscheinlichkeit für das Eintreten eines Ereignisses in einem fest definierten Zeitraum (kann auf die Erkrankung oder die Heilung bezogen werden).
- *Mortalität*: Wahrscheinlichkeit für das Eintreten des Tods.
- *Letalität*: Bedingte Wahrscheinlichkeit für das Eintreten des Tods bei Vorliegen einer bestimmten Erkrankung oder bei der Durchführung eines bestimmten diagnostischen oder therapeutischen Verfahrens.
- *Sensitivität*: Bedingte Wahrscheinlichkeit, dass ein diagnostischer Test bei einem Kranken positiv ist und die Krankheit anzeigt.
- *Spezifität*: Bedingte Wahrscheinlichkeit, dass ein diagnostischer Test bei einem Gesunden negativ ist und dessen Gesundsein anzeigt.
- *Positiver prädiktiver Wert (PPV)*: Bedingte Wahrscheinlichkeit, dass eine testpositive

Person tatsächlich krank ist.

- *Negativer prädiktiver Wert (NPV)*: Bedingte Wahrscheinlichkeit, dass eine testnegative Person tatsächlich gesund ist.

8. Medizinische und / oder ökonomische Bewertung der Konsequenzen.

Am Ende des Entscheidungsbaums mündet jeder Pfad in einen bestimmten Gesundheitszustand (Individuum) oder eine Maßzahl (Population), die die medizinischen Konsequenzen repräsentiert. Für die adäquate Bewertung dieser Konsequenzen ist eine geeignete Maßzahl auszuwählen, wie beispielsweise die Remissionsrate, die Fünf-Jahres-Überlebensrate, die Restlebenserwartung oder qualitätskorrigierte Lebenserwartung (Quality Adjusted Life Years, QALY). In unserem Beispiel wird der Tod mit 0 bewertet und das Überleben mit 1, was indirekt der Wahl der Überlebenswahrscheinlichkeit als zu maximierende Größe entspricht. Ferner sind bei gesundheitsökonomischen EA die verschiedenen eingesetzten Maßnahmen entlang eines Ereignispfades mit Kosten zu bewerten und entlang des jeweiligen Pfades aufzusummieren.

Punkt 7 und 8 gehören dem Bereich der Informationsgewinnung an. Sowohl die Parameter für die Ereigniswahrscheinlichkeiten als auch die Parameter der Konsequenzenbewertung können sowohl in Primäruntersuchungen (d.h. eigener Datenerhebung) als auch in Sekundäranalysen (d.h. Desk Research) gewonnen werden. Liegen mehrere Ergebnisse aus der wissenschaftlichen Literatur oder aus medizinischen bzw. gesundheitsökonomischen Datenbanken vor, können Metaanalysen eingesetzt werden, um Maßzahlen eines hohen Evidenzgrads zu erhalten. Falls keine Sekundärdaten vorliegen und keine Primärerhebung durchgeführt werden kann, muss auf Expertenbefragungen zurückgegriffen werden (z.B. Delphimethode).

9. Festhalten und explizite Formulierung der Annahmen.

Der aufgestellte Entscheidungsbaum und seine Parameter bilden ein Modell, das bestimmte Annahmen über die Wirklichkeit macht. Man unterscheidet strukturelle und numerische Annahmen. Strukturelle Annahmen betreffen den Aufbau des Entscheidungsbaums, numerische die verwendeten Daten. Ein Beispiel für eine numerische Annahme wäre die Vermutung, dass zwei verschiedene diagnostische Tests voneinander unabhängig sind, d.h., das Testergebnis des ersten beeinflusst nicht die Sensitivität und die Spezifität des zweiten. Alle Prämissen, die bei der Erstellung eines Ereignisbaums und bei der Belegung mit Wahrscheinlichkeiten und Bewertungen der Gesundheitszustände gemacht worden sind, sind explizit zu formulieren, um deutlich zu machen, unter welchen Bedingungen die Ergebnisse der EA gelten bzw. zu interpretieren sind.

10. Berechnung der Erwartungswerte (Ausmitteln und Zurückfalten).

- *Ausmitteln (Averaging Out)*: Für jede Strategie wird der mit den Pfadwahrscheinlichkeiten gewichtete Mittelwert der diskontierten Ergebnisparameter (z.B. Kosten und QALY in einer Kosten-Wirksamkeits-Analyse) berechnet.
- *Zurückfalten (Folding Back)*: Nach Eliminieren aller unterlegenen Strategien aus dem Entscheidungsbaum steht die Strategie mit dem maximalen Erwartungswert fest. Aus dem Vergleich der Erwartungswerte zweier Strategien lassen sich inkrementelle medizinische Effekte, inkrementelle Kosten und die Relation zwischen beiden errechnen. Dieses Ergebnis wird auch als Basisergebnis bezeichnet.

11. Sensitivitätsanalysen.

Um die Stabilität der Ergebnisse zu prüfen, werden Annahmen systematisch verändert und die Auswirkung dieser Veränderung auf das Ergebnis der EA betrachtet. Verschiedene Ansätze

und Verfahren der Sensitivitätsanalysen werden nochmals gesondert im Kapitel „Sensitivitätsanalysen“ behandelt.

12. Interpretation der Ergebnisse.

Unter Berücksichtigung der getroffenen Annahmen und der Datenqualität ist das Ergebnis innerhalb des medizinischen und des ökonomischen Problemkontexts zu bewerten. Sowohl Richtung als auch Größe des Einflusses zufälliger und systematischer Parameterschätzungen auf das Ergebnis sind zu diskutieren. Bei der Interpretation der Ergebnisse sind das jeweilige Gesundheitssystem, die Kosten-Effekt-Relationen anderer finanzierter Gesundheitsprogramme und die eingenommene Perspektive zu berücksichtigen. Schließlich sind Aussagen über die Generalisierbarkeit bzw. die Übertragbarkeit der Schlussfolgerungen zu machen.

3.4.2.4 Entscheidungsbaumverfahren

Im Folgenden werden die einzelnen Rechenschritte bei der Durchführung einer Analyse eines Entscheidungsbaums vorgeführt. Damit die Rechnungen vom Leser leicht nachvollzogen werden können, wird bewusst ein einfaches hypothetisches Beispiel mit runden Zahlen gewählt. Auch wenn in der Praxis durchgeführte Analysen meist mit sehr viel mehr Aufwand durchgeführt werden müssen, demonstriert das Rechenbeispiel die Grundprinzipien der EA, wie sie auch in der „wirklichen“ Forschung angewandt werden.

Es wird wieder auf das oben bereits dargestellte Beispiel Bezug genommen. Die notwendigen Parameter für die Analyse ergeben sich aus dem folgenden Text:

Aufgabe: Diagnostischer Test

Betrachtet seien die Patienten mit dem Symptom S, das zu einem schwachen Verdacht auf die Krankheit D führt. Bei der Krankheit D handelt es sich um eine ernsthafte Krankheit, bei der die Mortalität durch eine frühzeitige medikamentöse Behandlung reduziert werden kann.

Für Patienten, die an der Erkrankung D leiden, beträgt die Ein-Jahres-Überlebenswahrscheinlichkeit (1J-Survival) ohne Behandlung 80 % und mit Behandlung 90 %. Patienten, die das erste Jahr überleben, gelten als geheilt. Die komplette medikamentöse Behandlung dauert ein halbes Jahr und kostet insgesamt 10.000 DM. Zwei von hundert Patienten mit dem Symptom S haben tatsächlich die Krankheit D. Bei den behandelten Nichterkrankten liegt die Ein-Jahres-Mortalität aufgrund von Behandlungskomplikationen und Nebenwirkungen bei 1 %. Nach Beendigung der medikamentösen Therapie ist nicht mehr mit Nebenwirkungen zu rechnen. Ein Labortest zur Diagnose der Krankheit kostet 20 DM. Er besitzt eine Sensitivität von 90 % und eine Spezifität von 95 %. Die Durchführung des Tests ist mit einem sofortigen Letalitätsrisiko von 0,1% behaftet.

a) Zur Ermittlung der Strategie mit dem größten medizinischen Nutzen soll zunächst eine rein medizinisch orientierte EA für die drei Alternativen „Behandlung“, „Keine Behandlung“ und „Test“ bei Patienten mit dem Symptom S durchgeführt werden.

b) Anschließend soll eine Kosteneffektivitätsanalyse für diese drei Alternativen durchgeführt werden.

Die Fragestellung (Schritt 1), die durch diese Angaben impliziert wird, könnte man folgendermaßen formulieren: Welche der drei medizinischen Vorgehensweisen ist vorzuziehen, wenn das 1J-Survival bei den Patienten mit dem Symptom S maximiert werden soll:

Alle Patienten behandeln.

Keinen Patienten behandeln.

Erst den diagnostischen Test durchführen und dann nur die testpositiven Patienten behandeln.

Für die medizinische EA soll die Perspektive des Patienten eingenommen werden (Schritt 2). Das vorliegende Beispiel legt einen Zeithorizont von einem Jahr nahe, da laut Angaben alle relevanten Ereignisse innerhalb dieses Jahrs abgelaufen sind (Schritt 3). Die alternativen Strategien (Schritt 4) werden in der Fragestellung bereits explizit formuliert. Die Testalternative wird als Strategie mit einer „Wenn-dann-Regel“ angegeben. Als klinische Gesundheitszustände wer-

den Tod und Überleben ausgewählt, die durch das epidemiologische Maß der 1J-Survival repräsentiert werden (Schritt 5). Da wir keine Angaben über die Lebensqualität innerhalb des ersten Jahrs haben, der Zeitraum der Lebensqualitätseinschränkung aber im Vergleich der normalen Lebenserwartung selbst bei älteren Menschen relativ kurz ist, wird innerhalb des betrachteten Jahrs die Lebensqualitätseinschränkung durch die Medikation oder Krankheit in der Analyse nicht berücksichtigt. Das gesamte relevante Geschehen einschließlich der Konsequenzen findet innerhalb eines Jahrs statt. Auch die Kosten gehen nicht über diesen Zeitraum hinaus. Da also ein relativ kurzer Zeitraum betrachtet wird, bietet sich das Entscheidungsbaumverfahren als Analysemethode an.

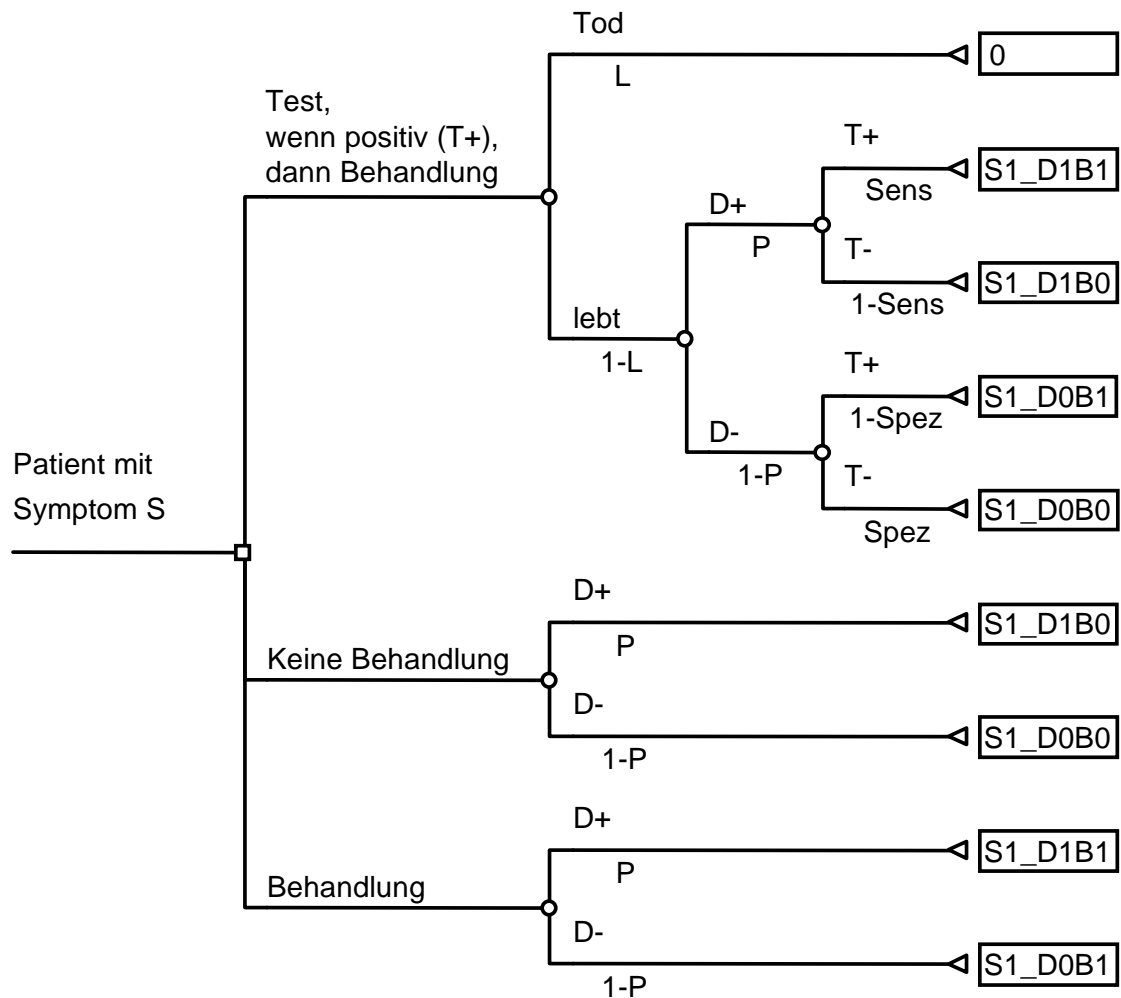
Der Entscheidungsbaum in Abbildung 3 repräsentiert ein einfaches Modell für das vorliegende Entscheidungsproblem und stellt den Ereignisablauf gemäß den gewählten Rahmenbedingungen dar (Schritt 6 bis 8). Die verwendeten Abkürzungen für die Gesundheitszustände, die Ereigniswahrscheinlichkeiten und Konsequenzen sind in Tabelle 1 erklärt.

Der Baum beginnt an seinem Stamm links in der Darstellung, wo die Zielpopulation des Entscheidungsproblems genannt wird. Es folgt ein so genannter Entscheidungsknoten (Rechteck), der die Wahl (Entscheidung) zwischen verschiedenen Handlungsalternativen repräsentiert. Anstelle des Begriffs Handlungsalternativen werden je nach Kontext oft folgende Begriffe synonym verwendet: Alternativen, medizinische Vorgehensweisen, medizinische Verfahren, Interventionen, medizinische Technologien, Procedere, Strategien, Optionen. Für jede Handlungsalternative ergeben sich Stellen im Baum, die durch so genannte Zufallsknoten (Kreise) gekennzeichnet sind. An einem Zufallsknoten (Synonym: Ereignisknoten) können verschiedene Ereignisse eintreten oder Merkmale offen gelegt werden, die nicht vorhersehbar sind und damit die Unsicherheit des Entscheidungsproblems verkörpern. Es gibt verschiedene Ereignispfade, die vom Stamm zur rechten Seite des Entscheidungsbaums führen und dort an den „Blättern“ des Baums, Ergebnisknoten enden. An den Ergebnisknoten stehen die interessierenden Konsequenzen der Entscheidung, Beispiele sind epidemiologische Maßzahlen, Gesundheitszustände, Laborwerte oder Kosten. Ein Ereignis mit allen nachfolgenden Ereignispfaden nennt man einen Ast des Ereignisbaums. Prinzipiell können auch innerhalb des Baums Entscheidungsknoten auftreten, in diesem Fall lässt sich allerdings der Baum entsprechend umstrukturieren, so dass es nur einen Entscheidungsknoten zu Beginn gibt. Die Regeln und Algorithmen der Entscheidungsbaumtransformation sind im Lehrbuch von Weinstein et al.⁷⁶ nachzulesen.

Der Entscheidungsbaum des Testbeispiels soll anhand Abbildung 3 erläutert werden. Für die Handlungsalternativen „Behandlung“ und „Keine Behandlung“ besteht jeweils nur Unsicherheit über das Vorliegen der Erkrankung (Prävalenz) und über das anschließende Überleben. Bei der Alternative „Test“ existiert Unsicherheit bezüglich des letalen Ausgangs des diagnostischen Eingriffs und über das Vorliegen der Krankheit. Ferner besteht Unsicherheit bei der Identifikation des Erkrankungsstatus durch den Test und schließlich wieder Unsicherheit bezüglich des Überlebens. Die Alternative „Test“ stellt eine Strategie im oben definierten Sinn dar. Die Vorschrift lautet „Durchführung des diagnostischen Tests; wenn Test positiv, dann behandeln, sonst nicht behandeln“. Es sei angemerkt, dass dies nicht unbedingt die beste Strategie zur Verwendung eines Tests sein muss. Zum Beispiel wird man diagnostische Tests mit einem hohen Letalitätsrisiko bei harmloseren Erkrankungen vermeiden. Aber selbst wenn der Test kein Letalitätsrisiko besitzt, kann es bei einer sehr geringen Prävalenz der Erkrankung und / oder einem sehr ungünstigen Wirksamkeits-Nebenwirkungs-Verhältnis der Erkrankung vorkommen, dass selbst nach einem positiven Testergebnis die Alternative „Keine Behandlung“ der

„Behandlung“ vorzuziehen ist. In diesem Fall wäre der Test sinnlos und die EA hätte das Resultat, dass in jedem Fall die Strategie „Keine Behandlung“ allen Strategien mit Testen überlegen wäre. In der Abbildung stellen die rechteckigen Knoten Entscheidungsknoten dar, die runden Knoten Zufallsknoten und die dreieckigen Knoten die Ergebnisknoten mit den medizinischen Konsequenzen, dem 1J-Survival. Mit D+ wird der Zustand „Krank“, mit D- der Zustand „Gesund“, mit T+ ein positives Testergebnis und mit T- ein negatives Testergebnis bezeichnet.

Abbildung 3: Entscheidungsbaum mit Variablen für das Beispiel Test / Behandlung / Keine Behandlung.



P = Prävalenz der Krankheit. D- = Zustand „gesund“. D+ = Zustand „krank“. L = Letalität des diagnostischen Tests. S1_D1B1 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Kranken mit Behandlung. S1_D1B0 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Kranken ohne Behandlung. S1_D0B1 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Gesunden mit Behandlung. S1_D0B0 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Gesunden ohne Behandlung.

Tabelle 1: Begriffe für Ereigniswahrscheinlichkeiten und Konsequenzen im Entscheidungsbaum.

Begriff	Variablenname im Baum	Symbol-schreibweise	Wert im Beispiel
Prävalenz der Krankheit	P	P(D+)	0,02
Sensitivität = Wahrscheinlichkeit eines positiven Testergebnisses unter den Kranken	Sens	P(T+ D+)	0,90
Spezifität = Wahrscheinlichkeit eines negativen Testergebnisses unter den Gesunden	Spez	P(T- D-)	0,95
Letalität des diagnostischen Tests	L	L	0,001
1J-Survival bei den Gesunden ohne Behandlung	S1_D0B0	S(D-, B-)	1,00
1J-Survival bei den Gesunden mit Behandlung	S1_D0B1	S(D-, B+)	0,99
1J-Survival bei den Kranken ohne Behandlung	S1_D1B0	S(D+, B-)	0,80
1J-Survival bei den Kranken mit Behandlung	S1_D1B1	S(D+, B+)	0,90
Kosten für den Test	K_T	K(T)	20 DM
Kosten für die Behandlung	K_B	K(B)	1.000 DM

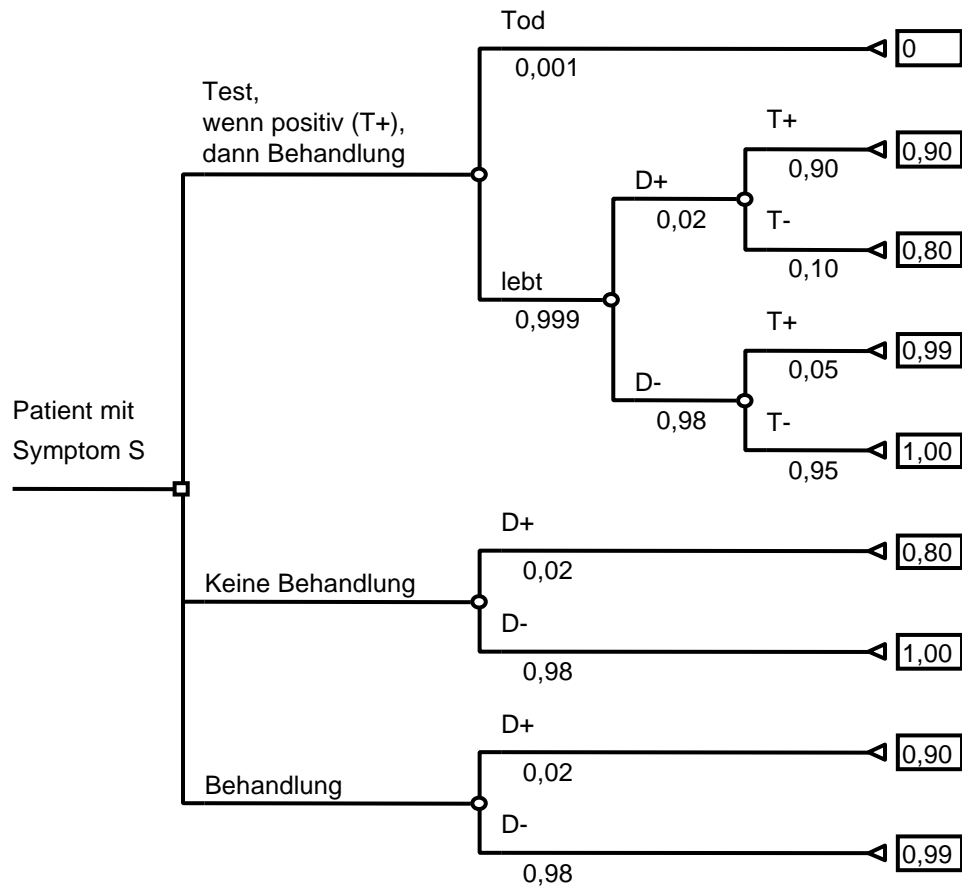
1J-Survival = Ein-Jahres-Überlebenswahrscheinlichkeit.

Es sei kurz auf einige Annahmen eingegangen (Schritt 9), die der verwendete Entscheidungsbaum impliziert. Die Tatsache, dass für den Ast „Test“ dieselben vier Überlebenswahrscheinlichkeiten den Blättern zugeordnet werden wie in den Ästen „Behandlung“ und „Keine Behandlung“, verdeutlicht die Annahme einer Unabhängigkeit der vier Überlebenswahrscheinlichkeiten vom Überleben des diagnostischen Tests. Würde beispielsweise der Test eher für die Kranken letal enden, dann wäre er als „Filter“ zu verstehen, der die Kranken „aussiebt“ und die Prävalenz der Krankheit wäre nach dem Test (bei den Überlebenden) geringer als die Prävalenz vor dem Test (gesamte Zielpopulation). Ebenso würden sich beispielsweise die Überlebenswahrscheinlichkeiten ändern, wenn eine Beziehung zwischen dem Behandlungserfolg und dem Überleben des Tests bestehen würde.

Im nächsten Schritt sind die Ereigniswahrscheinlichkeiten zu bestimmen (Schritt 10). Die in der Aufgabe angegebenen Parameter werden als Wahrscheinlichkeiten und bedingte Wahrscheinlichkeiten geschrieben.

Der Entscheidungsbaum (Abbildung 3) zeigt, dass die Parameter Prävalenz, Sensitivität und Spezifität die Ereignisse innerhalb des Entscheidungsbaums bestimmen, während die vier verschiedenen Überlebenswahrscheinlichkeiten an den Blättern des Baums die medizinischen Konsequenzen darstellen. In Abbildung 4 werden die Variablen im Entscheidungsbaum durch die aktuellen Werte des Beispiels ersetzt.

Abbildung 4: Entscheidungsbaum mit Ereignis- und Ein-Jahres-Überlebenswahrscheinlichkeit für das Beispiel Test / Behandlung / Keine Behandlung.



D- = Zustand „gesund“. D+ = Zustand „krank“.

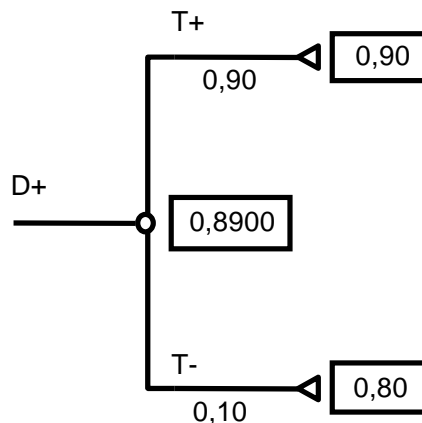
Der Erwartungswert für jede Alternative wird gebildet durch die gewichtete Summe aus den einzelnen Werten an den Blättern des Baums. Die Gewichte repräsentieren die Wahrscheinlichkeiten der einzelnen Pfade innerhalb einer Alternative. Sie ergeben sich durch Multiplikation aller Eintrittswahrscheinlichkeiten entlang eines Pfades. Für die praktische Berechnung der Erwartungswerte gibt es verschiedene Rechenansätze, die alle zum gleichen Ergebnis führen. Eine Möglichkeit sieht folgendermaßen aus:

$$\begin{aligned}
 &E(\text{Keine Behandlung}) \\
 &= P(D+) \cdot S(D+, B+) + P(D-) \cdot S(D-, B+) \\
 &= 0,020 \cdot 0,80 + 0,98 \cdot 1,00 = 0,9960 \\
 &E(\text{Behandlung}) \\
 &= P(D+) \cdot S(D+, B-) + P(D-) \cdot S(D-, B-) \\
 &= 0,020 \cdot 0,90 + 0,98 \cdot 0,99 = 0,9882 \\
 &E(\text{Test}) \\
 &= L \cdot 0 + (1-L) \cdot P(D+) \cdot P(T+|D+) \cdot S(D+, B+) + (1-L) \cdot P(D+) \cdot [1-P(T+|D+)] \cdot S(D+, B-) \\
 &\quad + (1-L) \cdot P(D-) \cdot [1-P(T-|D-)] \cdot S(D-, B+) + (1-L) \cdot P(D-) \cdot P(T-|D-) \cdot S(D-, B-) \\
 &= 0,001 \cdot 0 + 0,999 \cdot 0,02 \cdot 0,90 \cdot 0,90 + 0,999 \cdot 0,02 \cdot 0,10 \cdot 0,80 + 0,999 \cdot 0,98 \cdot 0,05 \cdot 0,99 + \\
 &0,999 \cdot 0,98 \cdot 0,95 \cdot 1,00 \approx 0,9963
 \end{aligned}$$

E = Erwartungswert. P = Prävalenz der Krankheit. D- = Zustand „gesund“. D+ = Zustand „krank“. B+ = Behandlung. B- = keine Behandlung. L = Letalität des diagnostischen Tests. S = Survival.

Eine weitere (effizientere) Art, die Analyse durchzuführen, besteht in einem Verfahren, das Ausmitteln und Zurückfalten genannt wird. Hierbei wird auf der rechten Seite des Baums begonnen und für die letzten (am weitesten rechts stattfindenden) Verzweigungen jeweils die Erwartungswerte für den jeweiligen gesamten Zufallsknoten berechnet. Abbildung 5 zeigt zum Beispiel im „Test“-Ast den Erwartungswert für den Zufallsknoten am Ende des Pfads „lebt und D+“, bevor das Testergebnis bekannt ist.

Abbildung 5: Erwartungswert an einem Knoten.



T+ = Positives Testergebnis. T- = Negatives Testergebnis. D+ = Zustand „Krank“.

Der Erwartungswert für diesen Zufallsknoten errechnet sich folgendermaßen:

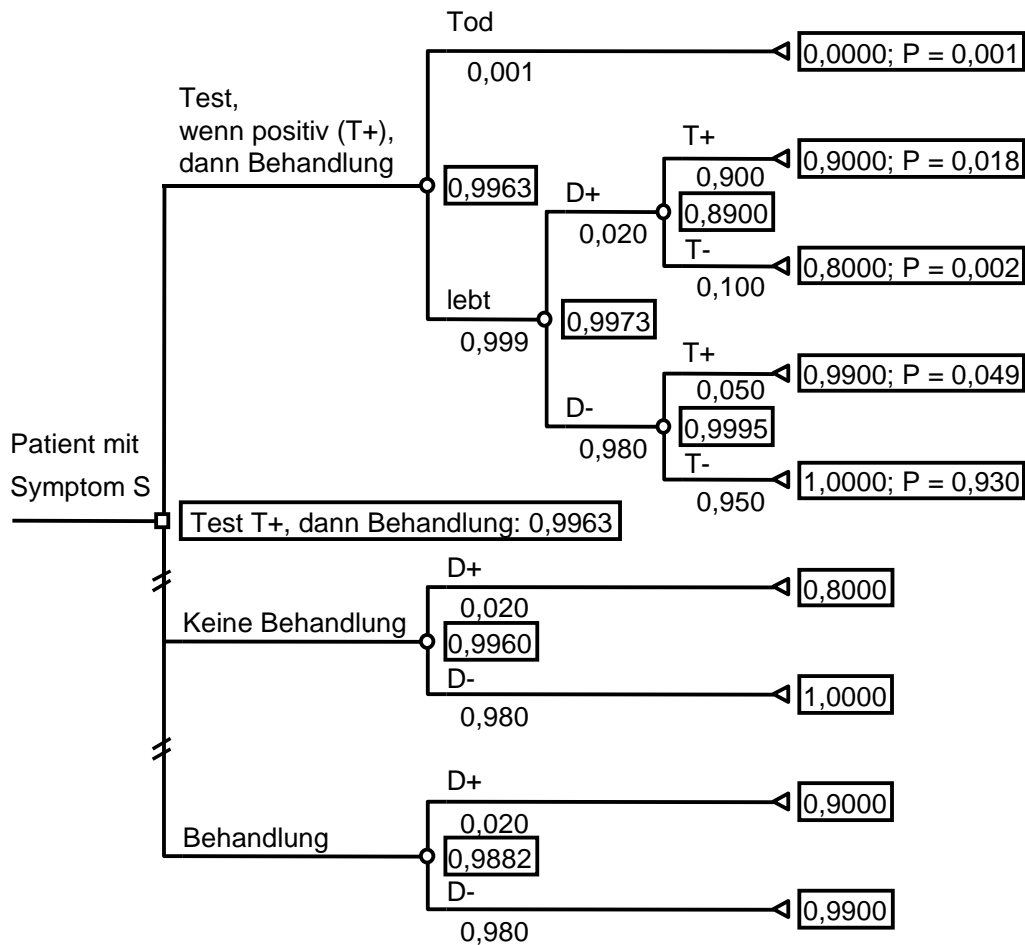
$$\begin{aligned}
 E(\text{Knoten}) &= P(T+|D+) \cdot S(D+, B+) + [1-P(T+|D+)] \cdot S(D+, B-) \\
 &= 0,90 \cdot 0,90 + 0,10 \cdot 0,80 = 0,8900
 \end{aligned}$$

E = Erwartungswert. P = Prävalenz der Krankheit. D+ = Zustand „Krank“. T+ = Positives Testergebnis. B+ = Behandlung. B- = keine Behandlung. S = Survival.

Der Begriff Zurückfalten stammt von der Vorstellung, dass man z.B. zwei Äste rechts von einem Knoten kollabiert bzw. zusammenfasst („zusammenfaltet“) und einen Erwartungswert für diesen Knoten berechnet. Auf diese Weise können von rechts nach links für alle Knoten des Entscheidungsbaums Erwartungswerte berechnet (Ausmitteln), und die entsprechenden nach rechts verzweigenden Äste des jeweiligen Knotens gelöscht werden (Zurückfalten). Dieses Vorgehen wird solange wiederholt, bis für jede Handlungsalternative ein Erwartungswert berechnet ist, der einen Vergleich der Handlungsalternativen ermöglicht.

Abbildung 6 zeigt den vollständig ausgemittelten und zurückgefalteten Baum (die zurückgefalteten Äste werden zur Visualisierung des Prozesses nicht gelöscht). Am Entscheidungsknoten werden zur Veranschaulichung der optimalen Entscheidung alle Handlungsalternativen gestrichen, die einen geringeren Erwartungswert als die optimale Strategie besitzen. Jedem Knoten wird ein Erwartungswert zugeordnet, der ein gewichtetes Mittel der nach rechts entspringenden Zweige darstellt. Das Vorgehen führt zu einem Erwartungswert für jede Handlungsalternative. Auf der rechten Seite sind zusätzlich die Wahrscheinlichkeiten für jeden Pfad innerhalb der jeweiligen Handlungsalternative angegeben.

Abbildung 6: Ausmitteln und Zurückfallen am Entscheidungsbaum.



P = Prävalenz der Krankheit. D- = Zustand „gesund“. D+ = Zustand „Krank“.

Die Ergebnisse des vorgestellten Beispiels sind folgendermaßen zu interpretieren (Schritt 12): Die Handlungsalternative „Behandlung“ hat die geringsten Aussichten. Erwartet wird eine Überlebenswahrscheinlichkeit von 98,82 % aller Patienten. Die Alternative „Keine Behandlung“ lässt ein 1J-Survival von 99,60 % erwarten. Am besten ist die Strategie „Test“ zu bewerten: Ihr erwartetes 1J-Survival beträgt 99,63 %. Da es bei Entscheidungen immer um den Vergleich von Optionen geht, sind die Differenzen zu berechnen und zu interpretieren. So lässt sich die inkrementelle Effektivität der „Test“-Strategie versus „Keine Behandlung“ errechnen als Differenz zwischen 99,63 und 99,60 %; dies ergibt ein zusätzliches 1J-Survival von 0,03 % durch die Einführung eines Tests anstelle der zweitbesten Strategie „Keine Behandlung“. Obwohl 0,1 % der Patienten durch den diagnostischen Test selbst sterben, ist das Nettoüberleben bei der „Test“-Strategie höher, was darauf zurückzuführen ist, dass weniger kranke Patienten unbehandelt sterben.

Man kann unschwer erkennen, dass das gefundene Ergebnis von bestimmten Parametern abhängt. Bei einer höheren Testletalität wäre die „Test“-Alternative unterlegen. Bei einer sehr hohen Prävalenz der Erkrankung kann es besser sein, gleich zu behandeln, da man dadurch das Testrisiko umgehen kann. Ist umgekehrt die Prävalenz wie im vorliegenden Fall sehr gering, so kann es medizinisch effektiver sein weder zu testen, noch zu behandeln, da bei einer Nichtbehandlung nur noch wenige tatsächlich Kranke negativ betroffen sind. Natürlich beein-

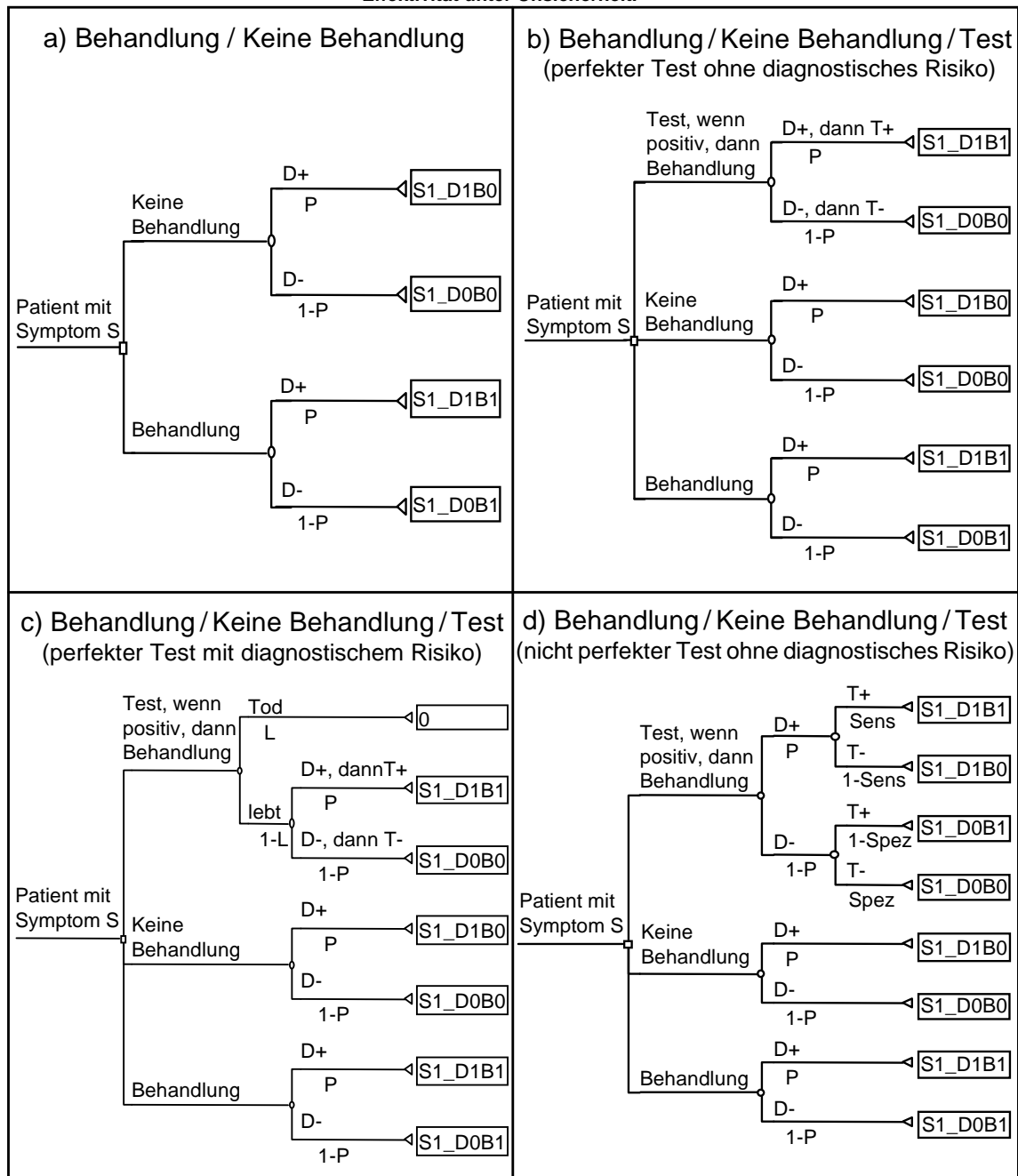
flussen nicht nur Änderungen der Ereigniswahrscheinlichkeiten, sondern auch Modifikationen der Konsequenzenparameter die Entscheidung. Ist die Krankheit harmlos in dem Sinn, dass unbehandelte Kranke eine Überlebenschance nahe 100 % haben, so kann es durchaus sein, dass durch die Behandlung und ihre Nebenwirkungen insgesamt medizinisch mehr Schaden angerichtet wird als Nutzen möglich ist. Um also die Stabilität bzw. die Robustheit des Ergebnisses zu prüfen, werden die Modellannahmen und -parameter systematisch variiert und analysiert, ob diese Variationen eine Änderung des Ergebnisses bzw. der Entscheidung zur Folge haben. Dieser Typ von Analysen wird Sensitivitätsanalyse genannt und schließt sich an die Analyse des Basisergebnisses an (Schritt 11). Auf die verschiedenen Verfahren der Sensitivitätsanalysen wird im gleichnamigen Kapitel eingegangen.

Bei der Interpretation des vorliegenden Ergebnisses ist es wichtig, auf die Annahme zu verweisen, dass es im dargestellten Beispiel keinen Unterschied macht, ob ein Patient (durch den diagnostischen Test) sofort stirbt oder (durch die Behandlungsnebenwirkungen) erst im Laufe des Jahrs. Will man den Zeitpunkt des Eintritts des Tods bei der EA berücksichtigen, wird man bei der Modellierung und Analyse meist zu Markov-Modellen anstelle von Entscheidungsbäumen greifen (Kapitel „Entscheidungsbaumverfahren“).

Ein wichtiger Aspekt bei der Bewertung eines diagnostischen Verfahrens ist der so genannte „Erwartungswert der klinischen Information“. Dieser entspricht der Differenz der Erwartungswerte einer Handlungsalternative mit Testinformation (hier: „Test“) und ohne Testinformation (hier: „Keine Behandlung“). Der so genannte Bruttoerwartungswert der klinischen Information ignoriert das Risiko durch den Test, der Wert beträgt im obigen Beispiel 0,13 %. Subtrahiert man davon das Risiko, das mit der Informationsgewinnung einhergeht (hier: Letalität 0,1 %), ergibt sich der Nettoerwartungswert der klinischen Information, in diesem Fall die genannten 0,03 %. Ob es unter begrenzten Ressourcen sinnvoll ist, den Test für alle Patienten einzuführen, d.h., ob der errechnete klinische Nutzen bei den anfallenden Kosten zu rechtfertigen ist, kann in gesundheitsökonomischen Analysen berechnet werden. Letztere beziehen die Kosten als Zielgrößen mit ein.

Grundsätzlich gibt es verschiedene generische Typen von Entscheidungsbäumen. Die wichtigsten Grundmuster sind in Abbildung 7 aufgeführt. Sie sind alle ein Spezialfall des komplexeren Typs „Behandlung“, „Keine Behandlung“ und „Test“ (nicht perfekter Test mit diagnostischem Risiko), dem auch das Rechenbeispiel angehört. Wenn auch Entscheidungsbäume in der Praxis häufig stärker verzweigt und komplexer aufgebaut sind als die hier vorgestellten, so stellen viele dieser Modelle doch Sequenzen oder Verschachtelungen der in Abbildung 7 vorgestellten Grundmuster dar.

Abbildung 7: Verschiedene generische Entscheidungsbaumtypen zur Modellierung von medizinischer Effektivität unter Unsicherheit.



P = Prävalenz der Krankheit. D- = Zustand „Gesund“. D+ = Zustand „Krank“. S1_D1B1 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Kranken mit Behandlung. S1_D1B0 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Kranken ohne Behandlung. S1_D0B1 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Gesunden mit Behandlung. S1D0B0 = Ein-Jahres-Überlebenswahrscheinlichkeit bei den Gesunden ohne Behandlung. T+ = Positives Testergebnis. T- = Negatives Testergebnis.

Der einfachste Typ eines Entscheidungsbaums ist die Modellierung der Frage nach Behandlung oder Nichtbehandlung (Abbildung 7a). Hier hängt die Entscheidung allein von der Prävalenz der Erkrankung und dem Verhältnis der Konsequenzen mit und ohne Therapie ab. Der einfachste Fall einer diagnostischen Strategie ist ein diagnostischer Test, der kein Risiko mit sich bringt und perfekt ist, d.h., dessen Sensitivität und Spezifität je 100 % betragen und der damit die

sichere Bestimmung des Krankheitsstatus erlaubt (Abbildung 7b). Bei einer Strategie mit perfektem Test und diagnostischem Risiko wird eine Entscheidung für den Test nur getroffen, wenn der durch die Information über den Krankheitszustand erzeugte zusätzliche Nutzen größer ist, als der durch das diagnostische Risiko verursachte medizinische Schaden (Abbildung 7c). Bei einem nicht perfekten Test ohne diagnostisches Risiko determinieren die Prävalenz der Erkrankung, die Behandlungskonsequenzen und die Testgüte den Entscheidungsprozess (Abbildung 7d).

Nachdem die medizinische Entscheidungsanalyse abgeschlossen ist, soll nun die ökonomische Evaluation stattfinden. Aufgabe b) des Beispiels besteht in der Durchführung einer Kostenwirksamkeitsanalyse für die drei angegebenen Handlungsalternativen „Behandlung“, „Keine Behandlung“ und „Test“. In einer Kostenwirksamkeitsanalyse wird das Verhältnis der inkrementellen Kosten von Handlungsalternativen und den korrespondierenden inkrementellen medizinischen Effekten der Handlungsalternativen untersucht. Das Ergebnis wird in Form eines Kosteneffektivitätsverhältnisses ausgedrückt. Ausführlich wird die Kosteneffektivitäts- bzw. Kostenwirksamkeitsanalyse im Kapitel „Kosteneffektivität medizinischer Verfahren“ dargestellt.

Die Berechnung der Kosten und der inkrementellen Kosten verschiedener Handlungsalternativen folgt demselben Schema wie die Evaluation der medizinischen Effekte. Auch hier gilt es, die einzelnen Schritte zur Aufstellung und Berechnung des entscheidungsanalytischen Modells zu beachten. Zunächst ist die ökonomische Fragestellung präzise zu formulieren bzw. der gesundheitsökonomische Studientyp festzulegen (Schritt 1). Verschiedene Perspektiven können zu verschiedenen zu berücksichtigenden Kostenarten führen (Schritt 2). Wenn sich Kosten im Laufe der Zeit verändern, ist dies bei der Wahl des Zeithorizonts zu berücksichtigen (Schritt 3). Insbesondere sind bei der Definition der Zustände solche zu unterscheiden, die bezüglich der Versorgungskosten differieren, auch wenn diese rein medizinisch keine verschiedenen Gesundheitszustände darstellen (Schritt 5). Beispielsweise ist es bei der Analyse der medikamentösen Behandlung der Alzheimer-Erkrankung von besonderer Bedeutung für die Kostenfrage, ob ein Patient in einem Pflegeheim versorgt wird oder zu Hause von den Angehörigen. Jeder medizinische Zustand, den die Patienten bezüglich einer neurologischen Skala einnehmen können, ist demnach in zwei kostenspezifische Versorgungszustände zu unterteilen. So verwenden Neumann u. a. in ihrer pharmakoökonomischen Analyse zur Therapie der Alzheimer-Erkrankung folgende sieben Gesundheitszustände⁵³.

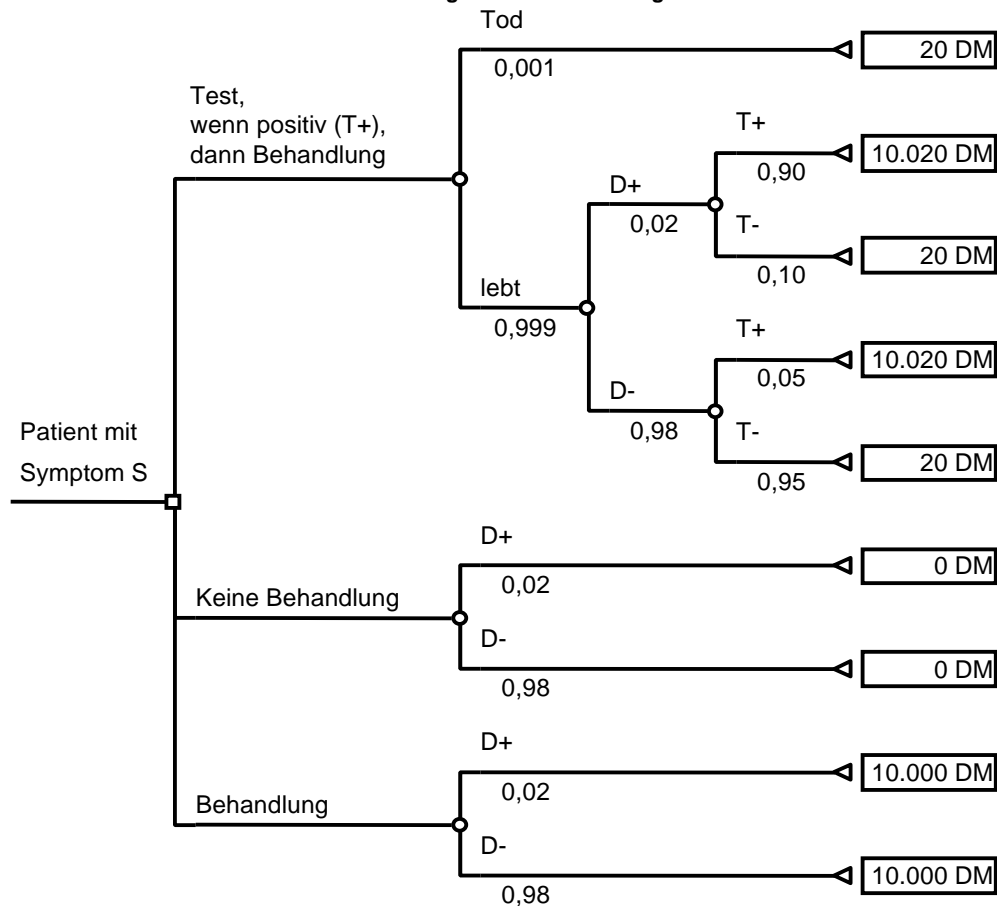
- Leichte Symptomatik / Pflegeheim.
- Leichte Symptomatik / nicht Pflegeheim.
- Mittlere Symptomatik / Pflegeheim.
- Mittlere Symptomatik / nicht Pflegeheim.
- Schwere Symptomatik / Pflegeheim.
- Schwere Symptomatik / nicht Pflegeheim.
- Tod.

Bei der Darstellung des Ereignisablaufs ist insbesondere auf den Übergang von ambulanter und stationärer Versorgung zu achten (Schritt 6 und 7). Schließlich stellen die Kosten selbst eine Form von Konsequenzen dar und werden deshalb als solche in den Entscheidungsbaum eingefügt (Schritt 8). Auch bei den Kosten sind Annahmen explizit zu formulieren (Schritt 9). Die Berechnung der erwarteten Kosten erfolgt wieder durch Ausmitteln und Zurückfalten (Schritt 10). Im Anschluss werden Sensitivitätsanalysen sowohl zu medizinischen Parametern als auch

zu Kostenparametern durchgeführt (Schritt 11). Bei der Interpretation der Ergebnisse ist vor allem die eingenommene Perspektive zu beachten und die Auswirkung der Berücksichtigung verschiedener Kostenarten wie z.B. direkte versus indirekte Kosten ist darzustellen (Schritt 12).

Abbildung 8 zeigt den Entscheidungsbaum des Testbeispiels und führt als Konsequenz die Kosten anstelle des 1J-Survivals auf. Es ergibt sich ein Kostenprofil, das aus vier Ausprägungen besteht: keine Kosten (0 DM), nur Testkosten (20 DM), nur Behandlungskosten (10.000 DM) sowie Test- und Behandlungskosten (10.020 DM).

Abbildung 8: Entscheidungsbaum mit Ereigniswahrscheinlichkeiten und Kosten für das Beispiel Test / Behandlung / Keine Behandlung.



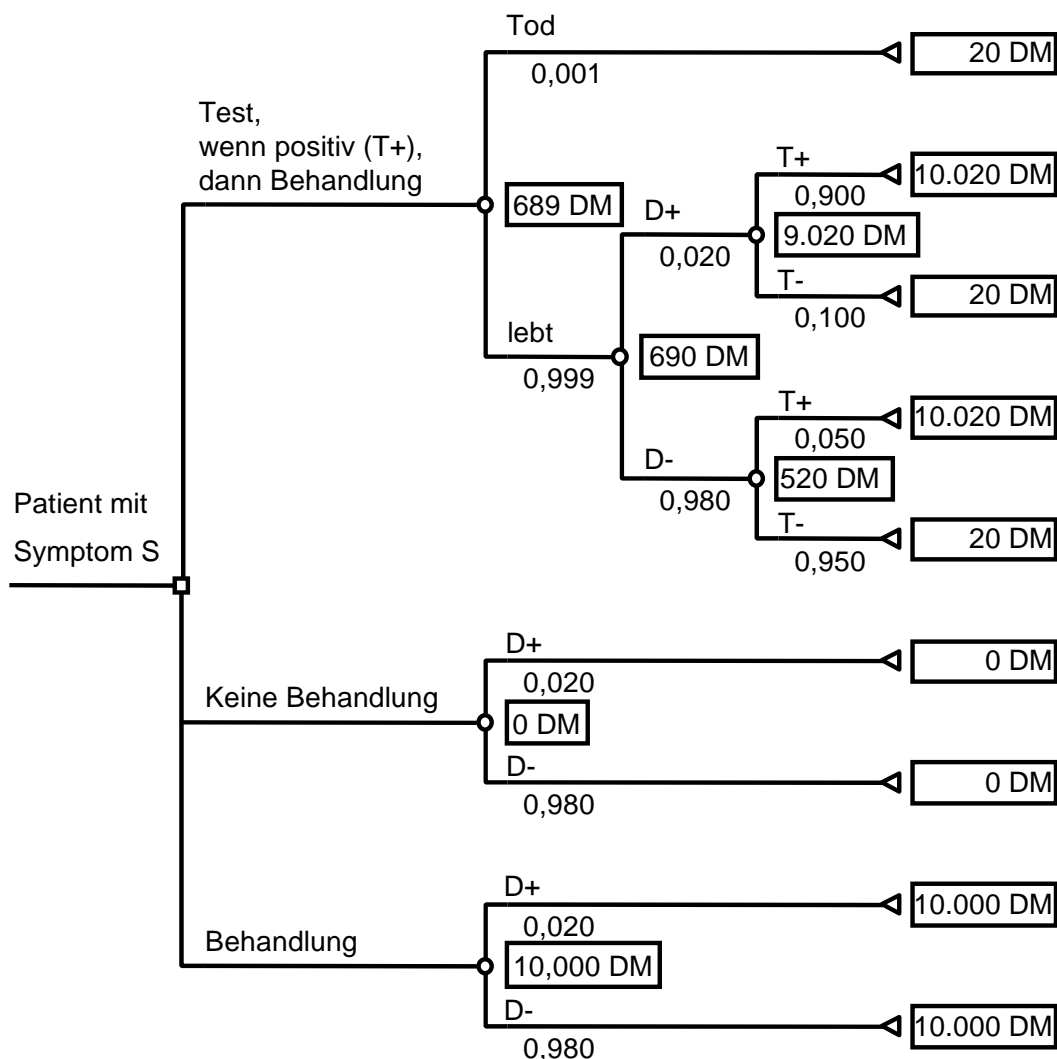
D- = Zustand „Gesund“. D+ = Zustand „Krank“.

Durch Ausmitteln und Zurückfalten ergeben sich folgende Erwartungswerte für die Kosten pro Patient bei den einzelnen Behandlungsstrategien:

E(Kosten von „Test“)	=	689,33 DM
E(Kosten von „Keine Behandlung“)	=	0,00 DM
E(Kosten von „Behandlung“)	=	10.000,00 DM

E = Erwartungswert.

Abbildung 9: Ausgemittelter und zurückgefalteter Entscheidungsbaum für die Kosten für das Beispiel Test / Behandlung / Keine Behandlung.



D- = Zustand „Gesund“. D+ = Zustand „Krank“.

Diese Werte sind ebenfalls aus dem ausgemittelten und zurückgefalteten Entscheidungsbaum für die Kosten abzulesen (Abbildung 9). Zunächst stellt sich der triviale Sachverhalt dar, dass die billigste Alternative darin besteht, niemanden zu behandeln (die monetäre Bewertung von durch Mortalität bedingter Produktivitätsausfallskosten wird hier der Einfachheit halber nicht vorgenommen). Da dies aber nicht die medizinisch optimale Alternative ist, sind hier Kosten und medizinischer Nutzen gegeneinander abzuwägen. Die Entscheidung kann nicht allein aufgrund der Kosten getroffen werden. Es wird allerdings deutlich, dass die Strategie „Test“ der Handlungsalternative „Behandlung“ sowohl medizinisch als auch ökonomisch überlegen ist, da sie bei weniger Kosten zu einem höheren 1J-Survival führt. Man bezeichnet diese Situation als Dominanz: die Alternative „Test“ dominiert die Alternative „Behandlung“. So bleibt schließlich nur noch die Entscheidung zwischen „Test“ und „Keine Behandlung“ zu treffen. Diese hängt unter utilitaristischen Kriterien nur davon ab, welchen Betrag man bereit ist, für ein gerettetes Lebensjahr auszugeben.

Im Folgenden wird der formale Prozess des Analysewegs wiedergegeben, wie er am Ende

einer EA steht, deren Ziel die Bestimmung des Kosteneffektivitätsverhältnisses ist (vgl. hierzu auch Kapitel „Kosteneffektivität medizinischer Verfahren“).

Tabelle 2: Kosteneffektivität der drei Handlungsalternativen im Beispiel.

(1)	(2)	(3)	(4)	(5)	(6)
Handlungs- alternativen	Kosten (in DM)	Inkrementelle Kosten (in DM)	1J- Survival	Inkremen- telles 1J- Survival	Inkrementelles Kosteneffektivitäts- verhältnis
	K	ΔK	S	ΔS	$\Delta K / \Delta S$
Keine Behandlung	0	-	0,9960	-	-
Test	689,33	689,33	0,9963	0,0003	2.204.516 DM / Leben
Behandlung	10.000	9310,76	0,9882	-0,0081	(Dominiert)

S = 1J-Survival = Ein-Jahres-Überlebenswahrscheinlichkeit. Δ = Absolut.

Die absoluten Kosten und die absoluten Werte für das 1J-Survival werden zuvor mit Hilfe des Entscheidungsbaumverfahrens ermittelt. Nun werden die drei Handlungsalternativen aufsteigend nach Kosten sortiert (Tabelle 2). Die drei Alternativen sind in Spalte 1 dargestellt, die dazugehörigen absoluten Kosten in Spalte 2 und die korrespondierenden 1J-Survival-Werte in Spalte 4. Anschließend werden die Inkremente für Kosten und Überleben gebildet: Von den absoluten Kosten einer Alternative werden die absoluten Kosten der nächstbilligeren Alternative subtrahiert, so dass man die inkrementellen Kosten einer Alternative erhält (Spalte 3). Analog wird vom absoluten 1J-Survival einer Alternative das absolute 1J-Survival der nächstbilligeren Alternative subtrahiert, so dass man das inkrementelle 1J-Survival erhält (Spalte 5). Schließlich ergibt sich als Quotient der inkrementellen Kosten und der inkrementellen medizinischen Effektivität das Kosteneffektivitätsverhältnis der jeweiligen Handlungsalternative (Spalte 6). Da im vorliegenden Fall die Alternative „Behandlung“ dominiert wird, erscheint an dieser Stelle kein Kosteneffektivitätsverhältnis und die Alternative muss aus der Analyse ausgeschlossen werden. Dies gilt auch für den Fall der so genannten „erweiterten Dominanz“. Eine Alternative ist erweitert dominiert, wenn die Alternative mit den nächsthöheren Kosten ein geringeres Kosteneffektivitätsverhältnis besitzt. Bei der Berechnung von Kosteneffektivitätsverhältnissen sollte sichergestellt werden, dass alle Inkremente nach Ausschluss der dominierten Alternativen und jeweils nach Ausschluss eventueller erweitert dominierten Alternativen neu berechnet werden. Diese Fragestellungen werden auch noch ausführlich im Kapitel „Kosteneffektivität medizinischer Verfahren“ dargestellt. Im Testbeispiel ergeben sich also zwei mögliche Alternativen, „Keine Behandlung“ und „Test“. Die Kosteneffektivität von „Test“ beträgt 2.204.516 DM pro gerettetem Leben. Geht man von einem Patienten mit einer Lebenserwartung von etwa vierzig Jahren aus, ergibt sich ein Wert von etwas über 50.000 DM pro gerettetem Lebensjahr. Legt man den in der Literatur²⁵ weitgehend akzeptierten Grenzwert von etwa \$50.000 pro gerettetem Lebensjahr zugrunde, wäre die „Test“-Strategie als kosteneffektiv zu bewerten und man würde sich bei der Zielpopulation für diese entscheiden.

Abschließend sollen noch einige Hinweise für die praktische Umsetzung einer Entscheidungsbaumanalyse gegeben werden. Ein Vorteil von Entscheidungsbäumen (auch im Vergleich zu Markov-Modellen) ist, dass sie sehr anschaulich die Alternativen darstellen und daher auch

unter Nichtfachleuten eine höhere Akzeptanz besitzen. Dies gilt allerdings nur für sehr einfache Entscheidungsbäume mit wenigen Ästen wie in dem dargestellten Beispiel. Entscheidungsbäume können jedoch komplex werden und es entstehen häufig Unterstrukturen im Baum, die sich häufiger wiederholen. In diesen Fällen ist der Entscheidungsbaum für Präsentationen oder Publikationen in Teilbäume zu zerlegen, um Vollständigkeit und Übersichtlichkeit zu gewährleisten. Ein Entscheidungsbaum aus etwa 150 Ästen, der komplett ausgedruckt etwa die Größe einer Haustür hat (was bei Präsentationen recht eindrucksvoll ist), kann in Teilen zerlegt oft auf drei bis fünf Seiten vollständig dargelegt werden (da sich einzelne dieser Teile im Baum häufiger wiederholen).

Bei der praktischen Erstellung eines Entscheidungsbaums gilt es immer, einen Kompromiss zu finden zwischen möglichst großer Strukturpräzision und möglichst großer Einfachheit, Transparenz und Übersichtlichkeit. Beteiligte Ärzte werden in der Regel darauf dringen, dass der Baum maximal komplex wird. Häufig hört man das Argument, dass jeder Patient anders ist und daher einen eigenen Ast benötigt. Ökonomen und Epidemiologen drängen dagegen häufig auf eine einfache Struktur, da nur dann ausreichend Daten zum „Füttern“ des Baums zur Verfügung stehen. Ebenso wie bei der Erstellung statistischer Modelle gilt der Grundsatz, die Struktur möglichst einfach und verständlich zu wählen, ohne dabei auf strukturelle Elemente zu verzichten, die für die Entscheidung relevant sein könnten. Ein Kompromiss ist in Abstimmung mit allen beteiligten Disziplinen zu finden. In der Regel wird man mit einfacheren Modellen beginnen. Erste Analysen zeigen dann bereits, für welche Elemente des Entscheidungsbaums strukturelle Erweiterungen sinnvoll und notwendig sind. Auf den Einsatz von Software bei EA wird weiter unten eingegangen.

3.4.2.5 Markov-Modelle

Neben der Entscheidungsbaumanalyse gehören die Markov-Modelle zu den mathematischen Verfahren der formalen EA. Ein Markov-Modell enthält eine endliche Zahl von disjunkten und erschöpfenden Gesundheitszuständen, die vom Patienten durchlaufen werden können. Die so genannten Übergangswege geben die möglichen Ereignisse im zeitlichen Verlauf an¹². Der zeitliche Verlauf wird in diskrete Zeitintervalle (Zyklen) eingeteilt, in denen die Übergangswahrscheinlichkeiten zwischen den einzelnen Gesundheitszuständen einzig und allein vom momentanen Gesundheitszustand abhängen. Markov-Modelle sind anzuwenden, wenn aus medizinischen oder aus ökonomischen Gründen längere Zeithorizonte zu wählen sind. Hierfür kann es drei Gründe geben⁶⁹:

1. Das Entscheidungsproblem beinhaltet zeitveränderliche Risiken, Lebensqualität oder Kosten.
2. Der Zeitpunkt des Eintretens eines bestimmten Ereignisses (Time-To-Event) spielt eine Rolle.
3. Relevante Ereignisse können mehrmals auftreten.
4. Beispielsweise muss bei der entscheidungsanalytischen Modellierung des Screenings für Prostatakarzinom^{4,39} auf Markov-Modelle zurückgegriffen werden, da
 - die Inzidenzraten mit fortschreitendem Lebensalter zunehmen.
 - der Unterschied zwischen dem Auftreten des Karzinoms in jungen Jahren und im hohen Alter mehrere Jahrzehnte betragen kann, so dass diese Ereignisse sowohl medizinisch als auch in Bezug auf einen potentiellen krankheitsbedingten Produktivitätsausfall unterschiedlich zu bewerten sind.

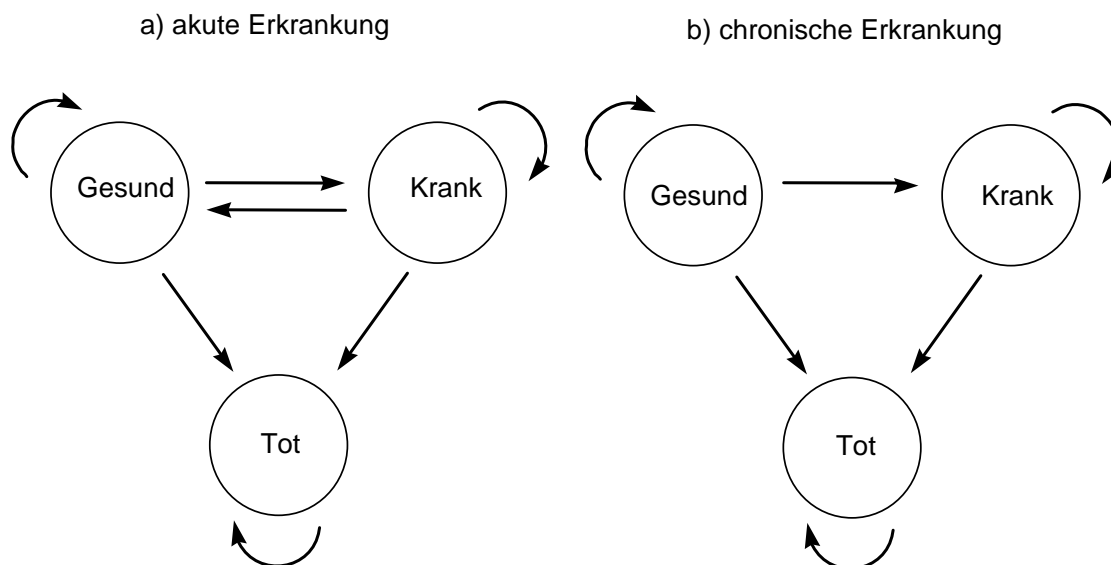
- falschpositive Ergebnisse erhöhte Kosten durch intensivierete Kontrolluntersuchungen in der Zukunft verursachen. Diese sind aufgrund der längeren Lebenserwartung bei jüngeren Screeningteilnehmern durchschnittlich höher als bei älteren.

Auch bei der Modellierung der Behandlung einer HIV-Infektion kommen Markov-Modelle zum Einsatz, da die Patienten um so länger in einem Zustand höherer gesundheitsbezogener Lebensqualität verweilen, je später ein Übergang ins AIDS-Stadium erfolgt. Außerdem können stationär behandlungsbedürftige Krankheitsereignisse im Langzeitverlauf mehrmals auftreten, so dass dies bei der Analyse der Kosten und der Lebensqualität zu beachten ist^{17, 18}.

Zur Berücksichtigung des zeitlichen Verlaufs wird die Zeit in aufeinander folgende Intervalle (Zyklen) aufgeteilt, innerhalb derer konstante Ereignisraten angenommen werden. Nach Ablauf eines Zyklus wechseln die Patienten vom aktuellen Gesundheitszustand in den Gesundheitszustand des nächsten Zyklus und verharren dort für eine Zyklusdauer etc. Die Erwartungswerte für Kosten und Effekte ergeben sich durch Summieren der zyklenspezifischen Werte⁷⁰. Die Schritte der Aufstellung eines Markov-Modells sowie seine Analyse ähneln denen des Entscheidungsbaumverfahrens, allerdings ist die Modellstruktur komplexer. Zusätzlich sind die möglichen Übergangswege mit den zugehörigen Übergangswahrscheinlichkeiten zwischen den einzelnen Gesundheitszuständen festzulegen. Jeder Gesundheitsstatus ist für jede Zyklusdauer mit Kosten und Effektmaß zu bewerten. Falls nötig, müssen Übergangswahrscheinlichkeiten, Kosten und Effekte für jeden Zyklus separat angesetzt werden. Eine Besonderheit des Markov-Modells ist die Annahme, dass die Übergangswahrscheinlichkeiten ausschließlich vom Gesundheitszustand des aktuellen Zyklus abhängen, d.h., das Markov-Modell hat kein „Gedächtnis“ für frühere Zustände. Sie wird Markov-Annahme genannt⁵.

Eine sehr anschauliche Darstellung eines Markov-Modells ist das so genannte Blasenendiagramm (Bubble Diagram) oder auch Erreichbarkeitsgraph genannt. Abbildung 10 zeigt je ein Blasenendiagramm für eine akute und eine chronische Erkrankung. Es sind verschiedene Übergänge zwischen den Gesundheitszuständen möglich. Den Tod nennt man einen absorbierenden Zustand, da kein Übergangsweg aus ihm herausführt. Es ist zu beachten, dass bei der chronischen Erkrankung in Abbildung 10b die Annahme zugrunde liegt, dass es keine Heilung der chronischen Erkrankung gibt, d.h., eine kranke Person kann nicht wieder in den Zustand „Gesund“ zurück.

Abbildung 10: Blasendiagramm zur Veranschaulichung eines Markov-Modells.



Dieses Blasendiagramm ist Grundlage für das folgende Rechenbeispiel:

Fallbeispiel 2: Prävention

Eine chronische Erkrankung K sei mit einem bestimmten Symptom S assoziiert, das für sich selbst genommen harmlos ist. Die Zeit zwischen dem Auftritt des Symptoms S und dem Ausbruch der Erkrankung K variiert von Patient zu Patient. Die Krankheit K ist mit einer reduzierten Lebensqualität und einer erhöhten Mortalität verbunden. Für Patienten mit dem Symptom S wird eine Prävention mit dem Arzneimittel A in Erwägung gezogen. Das Risiko des Auftretens der Krankheit K kann durch die dauerhafte Einnahme dieses Arzneimittels reduziert werden. Es wird allerdings festgestellt, dass das Arzneimittel A über einen unerwünschten Effekt zu einer direkten und sofortigen Erhöhung der Mortalität führt. Daraus folgt die Annahme, dass alle Risiken und Effekte über die Zeit hinweg konstant sind.

In Tabelle 3 und Tabelle 4 sind die so genannten Übergangsmatrizen für diese chronische Erkrankung dargestellt, die die Menge aller Übergangswahrscheinlichkeiten pro Zeitintervall (hier: 1 Jahr) zwischen den Gesundheitszuständen abbildet. Der Zustand „Gesund“ bezeichnet einen Patienten mit dem Symptom S, der noch nicht an der Krankheit K erkrankt ist. Die Ausgangszahlen zur Aufstellung der Übergangsmatrizen basieren auf folgenden Angaben:

Ohne Prävention: Die jährliche Sterbewahrscheinlichkeit der Patienten mit dem Symptom S wird mit 5 % angenommen. Die jährliche Krankheitswahrscheinlichkeit beträgt ohne Prävention 30 %. Für an der Krankheit K erkrankte Patienten ist die jährliche Sterbewahrscheinlichkeit 40 %. Die Krankheit ist chronisch, es gibt keine Genesung. Da bei Eintritt der Erkrankung die Einnahme des Arzneimittels keinen Sinn mehr macht, wird diese bei Erkrankungseintritt abgesetzt, so dass sich die Mortalität der erkrankten Patienten mit Prävention nicht von den erkrankten Patienten ohne Prävention unterscheidet.

Mit Prävention: Unter Prävention mit dem Arzneimittel A reduziert sich die jährliche Erkrankungswahrscheinlichkeit für die Erkrankung K auf ein Drittel (nun 10 statt 30 %). Allerdings erhöht sich die jährliche Sterbewahrscheinlichkeit bei Gesunden um 10 % (nun 15 statt 5 %).

Um die Lebensdauer mit der Lebensqualität zu gewichten, werden Nutzwerte eingesetzt.*

In unserem Beispiel ergeben sich aus einem „Time Trade-Off“-Verfahren folgende zeit-unabhängige Nutzwerte:

U1 = Nutzwert(Gesund) = 1

U2 = Nutzwert(Krank) = 0,6

U3 = Nutzwert(Tot) = 0

Frage: Ist eine Prävention mit dem Arzneimittel A aus medizinischer Sicht sinnvoll?

* Es gibt verschiedene Verfahren zur Bewertung von Gesundheitszuständen und Lebensqualität, Ihre Ausführung ginge jedoch über den Rahmen dieses Buchs hinaus. Das Kapitel von Böhmer und Kohlmann gibt eine Übersicht zu diesen Verfahren⁸.

Tabelle 3: Matrix der Übergangswahrscheinlichkeiten eines Markov-Modells für eine chronische Erkrankung mit den Gesundheitszuständen „Gesund“, „Krank“ und „Tot“ ohne Prävention.

Zustand	Gesund	Krank	Tot
Von Gesund nach ...	0,65	0,30	0,05
Von Krank nach ...	0,00	0,60	0,40
Von Tot nach ...	0,00	0,00	1,00

Tabelle 4: Matrix der Übergangswahrscheinlichkeiten eines Markov-Modells für eine chronische Erkrankung mit den Gesundheitszuständen „Gesund“, „Krank“ und „Tot“ mit Prävention.

Zustand	Gesund	Krank	Tot
Von Gesund nach ...	0,75	0,10	0,15
Von Krank nach ...	0,00	0,60	0,40
Von Tot nach ...	0,00	0,00	1,00

Anmerkung: In der Praxis stellt die Erhebung der Übergangswahrscheinlichkeiten aus Primärdaten oder der Literatur einen nicht unerheblichen Aufwand dar. Die Daten liegen meist in anderer Form vor, als sie für die Markov-Modelle benötigt werden. So sind z.B. Raten (Inzidenzdichten) in Wahrscheinlichkeiten (kumulative Inzidenzen) umzurechnen oder es müssen Fünf-Jahres- in Ein-Jahres-Wahrscheinlichkeiten umgerechnet werden. Während man bei Raten in diesem Fall problemlos durch 5 dividieren könnte, um das richtige Ergebnis zu erhalten, ist dies bei Wahrscheinlichkeiten aufgrund des sich innerhalb des Zeitraums verändernden Nenners (Risk Set) nicht ohne weiteres möglich. Detaillierte Beschreibungen der Umrechnungen mit den entsprechenden Formeln finden sich bei Miller und Horman⁴⁹. Im vorgestellten Fallbeispiel wird von bereits gegebenen bzw. errechneten Ein-Jahres-Wahrscheinlichkeiten ausgegangen. Ferner ist zu definieren, wann innerhalb eines Zyklus durchschnittlich das Auftreten eines Ereignisses (Genesung, Erkrankung, Tod) angenommen wird. Der Einfachheit halber wird in diesem Beispiel davon ausgegangen, dass alle Ereignisse am Ende eines Zyklus passieren. Es gibt andere Ansätze, z.B. dass ein Ereignis am Anfang bzw. in der Mitte des Zyklus eintritt (so genannte „Half-Cycle Correction“). Je kürzer die Zyklendauer des Markov-Modells gewählt wird, desto weniger beeinflussen diese Annahmen das Ergebnis.

Fügt man die gegebenen Werte in das entsprechende Blasendiagramm aus Abbildung 10 ein, ergibt sich das in Abbildung 11 dargestellte Modell. U1 bis U3 sind die Nutzwerte für die entsprechenden Gesundheitszustände. In diesem Beispiel werden die Übergangsraten und -wahrscheinlichkeiten als über die Zeit hinweg konstant angenommen. Eine weitere Veranschaulichung dieser Zahlen, die bereits in die mathematische Analyse übergeht, findet sich in Abbildung 12. Es handelt sich um die ersten zwei Zyklen der genannten chronischen Erkrankung. Es ist zu beachten, dass aufgrund der Chronizität der Krankheit der Übergangsweg von „Krank“ nach „Gesund“ nicht existiert.

Abbildung 11: Blasendiagramm mit Übergangswahrscheinlichkeiten und Nutzwerten (U1 bis U3) ohne und mit Prävention.

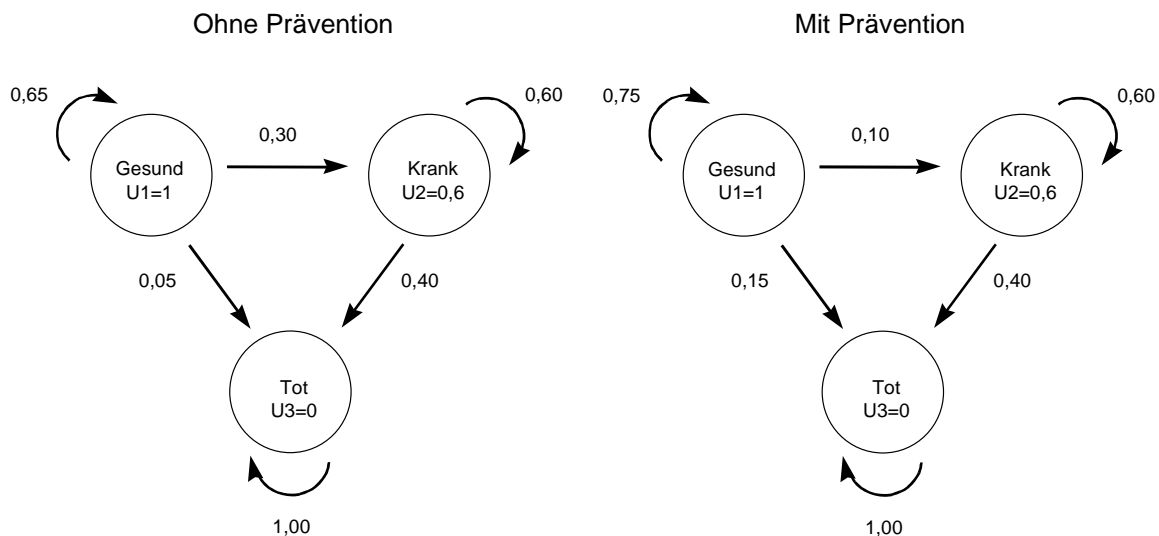
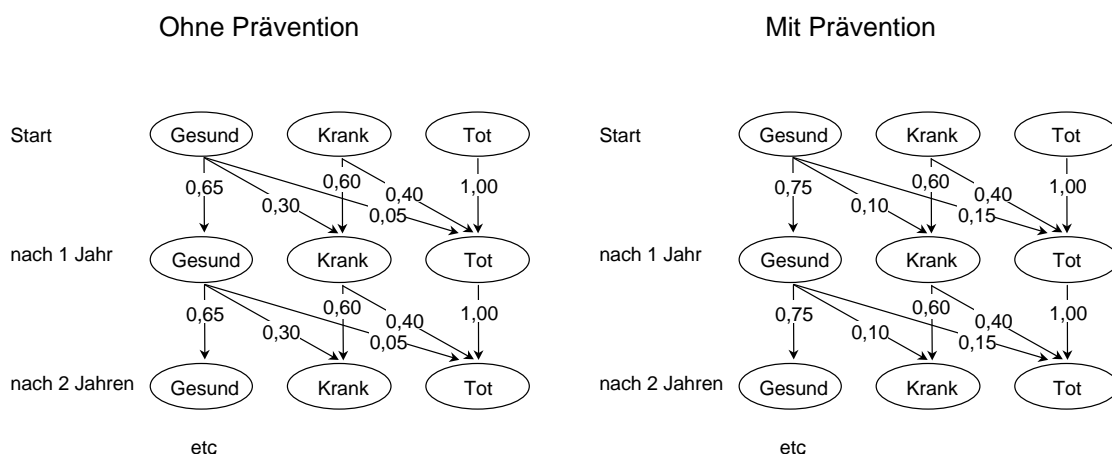


Abbildung 12: Schematische Darstellung eines Markov-Modells für eine chronische Erkrankung mit Gesundheitszuständen, Übergangswegen und Übergangswahrscheinlichkeiten für zwei Zyklen ohne und mit Prävention.



Es sei an dieser Stelle nochmals darauf hingewiesen, dass bei der vorliegenden Fragestellung die Zeit, die innerhalb der einzelnen Gesundheitszustände verbracht wird eine Rolle spielt, da sie mit den Nutzwerten (Utilities) gewichtet wird. Diese Zeit würde im Entscheidungsverfahren nicht berücksichtigt, sondern es würde nur der nach einem bestimmten Zeitraum bestehende Gesundheitszustand in das Ergebnis eingehen.

Um den Verlauf der Zustandswahrscheinlichkeiten über eine bestimmte Anzahl von Zyklen hinweg darzustellen wird eine so genannte Markov-Spur erstellt. Diese gibt für jeden Zyklus die erwartete Verteilung der Zustände an, so dass daraus die kumulative Zeit berechnet werden kann, die ein Patient für einen gegebenen Gesamtzeitraum erwartungsgemäß in den Zuständen „Gesund“, „Krank“ und „Tot“ verbringt. Diese Zeiten können dann wiederum mit den Nutzwerten gewichtet werden und ergeben als Summe die erwarteten QALY für jede Handlungsalternative. Die Markov-Spur ist in Tabelle 5 beispielhaft für die Strategie „Ohne Prävention“ dargestellt. Es sei dem Leser überlassen, die Markov-Spur für die Möglichkeit „Mit Prävention“ aufzustellen.

Die Lösung wird für beide Alternativen angegeben.

Tabelle 5: Markov-Spur für die Handlungsalternative „Ohne Prävention“ für 20 Zyklen (0-19).

Zyklus z	Gesund U1 = 1 Pz(G)	Krank U2 = 0,6 Pz(K)	Tot U3 = 0 Pz(T)	QALY pro Zyklus	Kumulative QALY
0	1,000	0,000	0,000	1,000	1,000
1	0,650	0,300	0,050	0,830	1,830
2	0,423	0,375	0,203	0,648	2,478
3	0,275	0,352	0,374	0,486	2,963
4	0,179	0,293	0,528	0,355	3,318
5	0,116	0,230	0,654	0,254	3,572
6	0,075	0,173	0,752	0,179	3,751
7	0,049	0,126	0,825	0,125	3,875
8	0,032	0,090	0,878	0,086	3,961
9	0,021	0,064	0,915	0,059	4,020
10	0,013	0,044	0,942	0,040	4,060
11	0,009	0,031	0,961	0,027	4,088
12	0,006	0,021	0,973	0,018	4,106
13	0,004	0,014	0,982	0,012	4,118
14	0,002	0,010	0,988	0,008	4,127
15	0,002	0,007	0,992	0,005	4,132
16	0,001	0,004	0,995	0,004	4,136
17	0,001	0,003	0,996	0,002	4,138
18	0,000	0,002	0,998	0,002	4,140
19	0,000	0,001	0,998	0,001	4,141
Summe „Ohne Prävention“	2,857	2,140	15,003	4,141	
Zum Vergleich: Summe „Mit Prävention“	3,987	0,992	15,021	4,582	

Die Ergebnisse sind für beide Markov-Spuren (ohne und mit Prävention) ohne Diskontierung dargestellt. Die angegebenen Ergebnisse sind gerundet, die Rechnungen erfolgen mit den exakten Werten. P_z(G), P_z(K) und P_z(T) stellen die Zustandswahrscheinlichkeiten für den Zyklus z dar. U1 bis U3 sind die entsprechenden Nutzwerte. P = Wahrscheinlichkeit für einen Gesundheitszustand. G = Gesund. K = Krank. T = Tot. QALY = Qualitätskorrigierte Lebensjahre.

Zur Nachvollziehbarkeit mittels Taschenrechner wird die Markov-Spur im Beispiel nach 20 Zyklen (0 bis 19) abgebrochen. Danach findet kein wesentlicher Beitrag zu den kumulativen QALY mehr statt. In diesem rein klinischen Beispiel wird keine Diskontierung vorgenommen.

Gemäß Abbildung 11 und Abbildung 12 lautet die allgemeine Rechenvorschrift zur Berechnung der Markov-Spur in diesem Beispiel für die Alternative „Ohne Prävention“:

$$\begin{aligned}
 P_{z+1}(G) &= P_z(G) \times 0,65 \\
 P_{z+1}(K) &= P_z(G) \times 0,30 + P_z(K) \times 0,60 \\
 P_{z+1}(T) &= P_z(G) \times 0,05 + P_z(K) \times 0,40 + P_z(T) \times 1
 \end{aligned}$$

P_z = Entsprechende Zustandswahrscheinlichkeiten.

Die erste Zeile stellt für die Alternative „Ohne Prävention“ die Verteilung der drei Zustände vor dem ersten Zyklus dar (Zyklus 0). Vor dem Zyklus 0 sind noch alle Personen gesund (1,000) und niemand krank (0,000) oder tot (0,000). Die Summe ergibt die QALY für diesen Zyklus und beträgt 1,000.

Die Verteilung nach dem ersten Zyklus ist in der zweiten Zeile (Zyklus 1) dargestellt. Sie ergibt sich aus den Zustandswahrscheinlichkeiten der ersten Zeile und den Übergangswahrscheinlichkeiten aus Tabelle 3 bzw. gemäß Abbildung 11 und Abbildung 12 (ohne Prävention):

$$P_1(G) = P_0(G) \times 0,65 = 1 \times 0,65 = 0,65$$

$$P1(K) = P0(G) \times 0,30 + P0(K) \times 0,60 = 1 \times 0,30 + 0 = 0,30$$
$$P1(T) = P0(G) \times 0,05 + P0(K) \times 0,40 + P0(T) \times 1 = 1 \times 0,05 + 0 + 0 = 0,05$$

Die Summe aus den mit den Übergangswahrscheinlichkeiten gewichteten Nutzwerten ergibt die QALY für diesen Zyklus:

$$0,65 \times 1 + 0,30 \times 0,6 + 0,05 \times 0 = 0,83$$

Addiert man diese Zahl zu den QALY aus dem ersten Zyklus, ergibt sich eine kumulative Summe von 1,83 QALY.

Für die dritte Zeile der Markov-Spur (Zyklus 2) folgt:

$$P2(G) = P1(G) \times 0,65 = 0,65 \times 0,65 = 0,4225$$
$$P2(K) = P1(G) \times 0,30 + P1(K) \times 0,60 = 0,65 \times 0,30 + 0,30 \times 0,60 = 0,375$$
$$P2(T) = P1(G) \times 0,05 + P1(K) \times 0,40 + P1(T) \times 1 = 0,65 \times 0,05 + 0,30 \times 0,40 + 0,05 \times 1 = 0,2025$$

Daraus ergibt sich eine Summe an QALY für den Zyklus 2 von:

$$0,4225 \times 1 + 0,375 \times 0,6 + 0,2025 \times 0 = 0,6475$$

Die kumulative Summe der QALY beläuft sich auf 2,4775 für die ersten drei Zyklen.

Zu beachten ist, dass in Tabelle 5 die Werte auf drei Dezimalen gerundet dargestellt sind, jedoch immer mit dem exakten Ergebnis weitergerechnet wird. Das Verfahren wird entsprechend bis zum Erreichen des letzten Zyklus fortgesetzt. Das Ergebnis der Markov-Spuren zeigt die Überlegenheit der Handlungsalternative „Mit Prävention“. Während die Alternative „Ohne Prävention“ nach 20 Zyklen einen Erwartungswert von 4,141 zu erwartenden QALY besitzt, ist der entsprechende Wert für die Alternative „Mit Prävention“ 4,582. Dies entspricht einer inkrementellen Effektivität von 0,441 QALY, die der Prävention zuzuschreiben ist. Die Summen unter den einzelnen Zuständen „Gesund“, „Krank“ und „Tot“ entsprechen den erwarteten kumulativen „Aufenthaltszeiten“ in jedem Gesundheitszustand. So ist für eine Population unter der Alternative „Ohne Prävention“ zu erwarten, dass im Mittel jedes Individuum 2,857 Jahre gesund ist und 2,140 Jahre krank, bevor der Tod eintritt (der Rest der 20 Jahre, d.h. 15,003 Jahre werden im Zustand „Tot“ verbracht). Man erkennt deutlich das Profil, das durch die Prävention erzeugt wird. Die Lebenserwartung ist bei beiden Alternativen etwa gleich, jedoch werden unter der Prävention mehr Jahre gesund statt krank verlebt. Dies veranschaulicht die Wirkung der Prävention auf die qualitätskorrigierten Lebensjahre.

Es sei angemerkt, dass die inkrementelle Effektivität von der Bewertung des Gesundheitszustands „Krank“ mit einem Nutzwert von 0,6 abhängt. Bei einem höheren Nutzwert für diesen Zustand nimmt die inkrementelle Effektivität kontinuierlich ab, da ohne Prävention die kumulative Zeit im Zustand Tod etwas niedriger ist als mit Prävention. Es wird damit deutlich, dass eine reine Überlebenszeitanalyse ein anderes Ergebnis erbracht hätte, als die durchgeführte Markov-Modellierung, die die Nutzwerte für die gesundheitsbezogene Lebensqualität für jeden Zyklus mitberücksichtigt. Die reine Lebenserwartung lässt sich einfach als Summe der kumulativen Zeiten in den Zuständen „Gesund“ und „Krank“ errechnen und ergibt mit 4,997 Jahren ohne Prävention einen etwas höheren Wert als für die Alternative mit Prävention, bei der die Lebenserwartung 4,979 Jahre beträgt. Die Markov-Spuren verdeutlichen anschaulich, dass eine reine Überlebenszeitanalyse der Zuordnung von Nutzwerten von Eins für die Zustände „Gesund“ und „Krank“ und der Zuordnung des Nutzwerts Null für den Zustand „Tod“ entspricht und damit die gesundheitsbezogene Lebensqualität unberücksichtigt bleibt.

Man unterscheidet bei Markov-Modellen zwischen Markov-Ketten und Markov-Prozessen. Ein Markov-Prozess heißt Markov-Kette, wenn der Parameterraum höchstens abzählbar und der

Zustandsraum endlich ist. Eine Markov-Kette heißt homogene Markov-Kette, wenn die Übergangswahrscheinlichkeiten für feste Zeitabstände konstant, d.h. unabhängig vom Zeitpunkt des Übergangs sind. Abbildung 12 zeigt eine Markov-Kette. Der Grundgedanke aller Typen von Markov-Modellen ist, dass für jede Handlungsalternative das in Abbildung 12 dargestellte Rechenverfahren durchgeführt wird und eine hypothetische Kohorte von Patienten die beschriebenen Zyklen durchläuft. In einer etwas genaueren Analyse müsste man im obigen Präventionsbeispiel etwa die sich mit zunehmendem Alter erhöhende Mortalität berücksichtigen. Damit würde sich für jeden Zyklus eine andere Matrix der Übergangswahrscheinlichkeiten ergeben. Ansonsten bliebe der Berechnungsmodus derselbe.

Bei gesundheitsökonomischen Evaluationen werden die Zeiten und die Kosten, die kumulativ in jedem der Gesundheitszustände anfallen, aufsummiert. Anschließend kann dieses Ergebnis deskriptiv zwischen den Handlungsalternativen verglichen oder zuvor auf eine einheitliche Skala wie QALY und / oder Kosten gebracht werden. Zur Ermittlung von QALY wird jeder in einem bestimmten Zustand verbrachte Zeitraum mit der entsprechenden gesundheitsbezogenen Lebensqualität gewichtet. Ferner können die in der Zukunft anfallenden QALY und Kosten diskontiert werden.

Damit entspricht das Vorgehen innerhalb einer Entscheidungsanalyse mit Markov-Modellen der bereits oben vorgestellten Schrittfolge (s. Kapitel „Design und Schritte der Entscheidungsanalyse“). Auch innerhalb der Markov-Modellierung werden für jede Handlungsalternative die erwarteten Konsequenzen errechnet und gegenübergestellt. In einer Kosten-Nutzwert-Analyse liefert das Markov-Modell sowohl die inkrementellen Kosten als auch den inkrementellen medizinischen Nutzwert, so dass das entsprechende Kosten-Nutzwert-Verhältnis (z.B. DM / QALY) bestimmt werden kann.

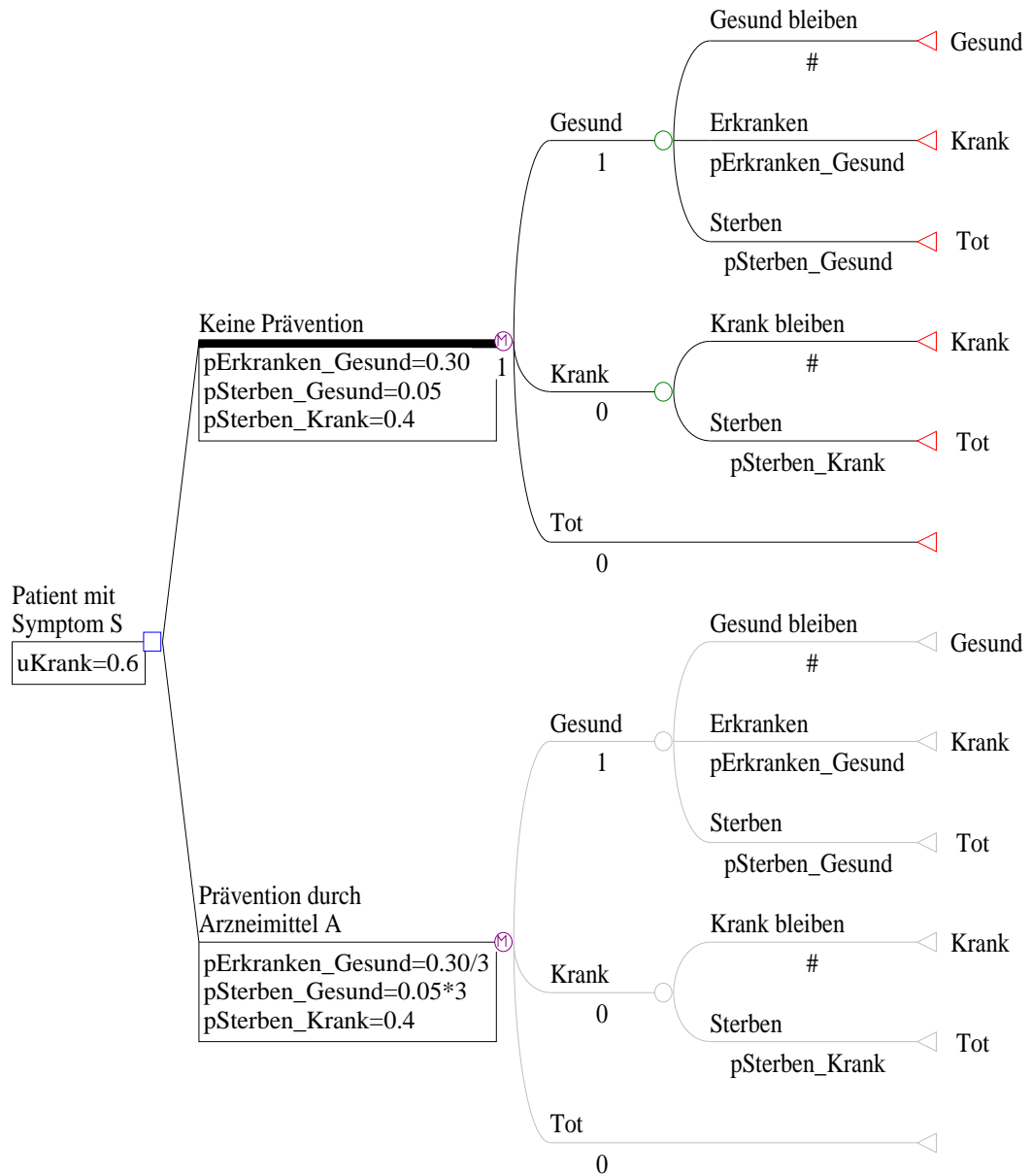
Für die Leser, die sich bereits mit der Software zur Darstellung von Markov-Modellen beschäftigt haben, wird in Abbildung 13 das visualisierte Markov-Modell für das Fallbeispiel Prävention wiedergegeben. An technischen Details weniger interessierte Leser können ohne inhaltlichen Bruch direkt zum Kapitel „Gesundheitsökonomische Entscheidungsanalysen“ springen.

Für die in Abbildung 13 dargestellte Analyse wird das Softwarepaket TreeAge Professional for Health Care (Fa. TreeAge, Williamstown, MA, USA) verwendet. In der Darstellung sind die knotenspezifischen Parameter mit angegeben. Parameter, die mit u beginnen stehen für Nutzwerte, so bedeutet z.B. uKrank = Nutzwert(Krank). Parameter, die mit p beginnen stehen für Wahrscheinlichkeiten bzw. bedingte Wahrscheinlichkeiten, so bedeutet pErkranken_Gesund die bedingte Wahrscheinlichkeit zu Erkranken unter den Gesunden, also $P(\text{Erkranken}|\text{Gesund})$. Die Modellstruktur ist dieselbe für beide Handlungsalternativen, die Software verwendet für die Alternative „Mit Prävention“ einen so genannten Klon der Alternative „Keine Prävention“. Ein Klon ist eine dynamische Kopie. Dynamisch bedeutet in diesem Zusammenhang, dass sich jede Programmänderung, die an der Klonvorlage vorgenommen wird, automatisch auf den Klon überträgt. Dies ist eine wesentliche Arbeitserleichterung und hilft dem Programmierer, Fehler zu vermeiden. Die numerischen Parameter für die einzelnen Äste können für jeden Ast individuell gesetzt werden. So gilt für das gesamte Modell der Nutzwert 0,6 für den Zustand „Krank“. Jedoch werden für die Äste „Keine Prävention“ und „Mit Prävention“ zum Teil verschiedene Parameter für die Übergangswahrscheinlichkeiten gewählt. Nach der jeweils ersten Gabelung in jeder Strategie (am so genannten Markov-Knoten) werden die möglichen Gesundheitszustände aufgeführt, also „Gesund“, „Krank“ und „Tot“. Die zugehörigen Wahrscheinlichkeiten repräsentieren die Startverteilung, in diesem Fall sind zu Beginn der Analyse alle Individuen gesund.

Die zweite Gabelung am Zufallsknoten repräsentiert die Übergangsmatrix und die zugehörigen Übergangswahrscheinlichkeiten (ein # bedeutet, dass an dieser Stelle die Wahrscheinlichkeitsverteilung automatisch zu 1 ergänzt wird). So gibt es zum Beispiel innerhalb des Zustands „Krank“ nur die Ereignisse „Krank bleiben“ oder „Sterben“, die dann zu den Ausgangszuständen des nächsten Zyklus führen, nämlich „Krank“ und „Tot“. Die Ausgangszustände für den nächsten Zyklus sind durch die dreieckigen Endknoten dargestellt. Die Anzahl der Zyklen ist nicht dargestellt, im Beispiel werden 20 Zyklen gewählt.

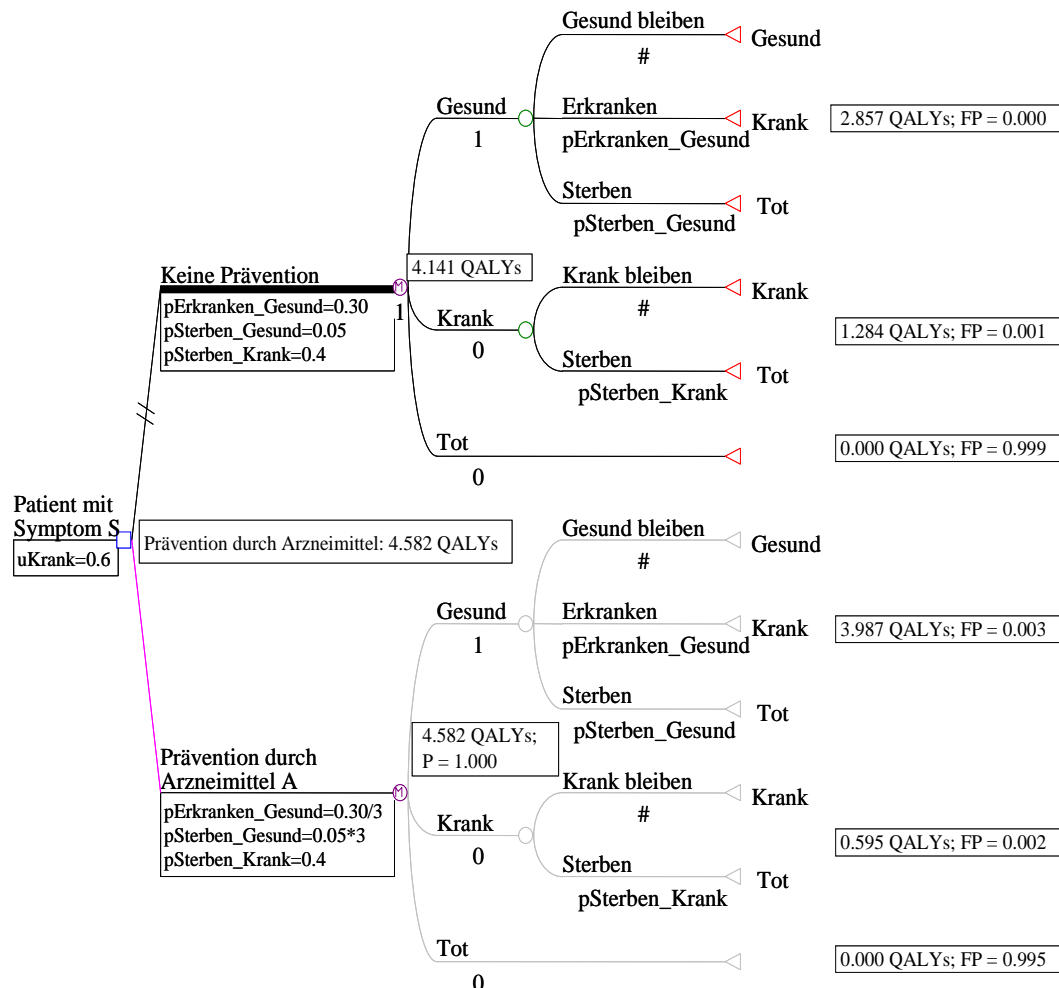
Abbildung 14 zeigt das Ergebnis der rechnerischen Analyse des Markov-Modells mit Angabe der erwarteten QALY je Handlungsalternative für einen Zeitraum von 20 Jahreszyklen. Für jeden Knoten werden die Erwartungswerte angegeben. Die Handlungsalternative mit dem größten aller Erwartungswerte wird als die präferierte Alternative durch Streichen aller anderen Äste gekennzeichnet. Das Ergebnis der bevorzugten Alternative wird am Entscheidungsknoten ganz links im Baum angegeben. In diesem Beispiel beträgt es 4,582 QALY. Die inkrementelle Effektivität kann als Differenz zwischen den Erwartungswerten der beiden Handlungsalternativen errechnet werden und beträgt im Beispiel $4,582 \text{ QALY} - 4,141 \text{ QALY} = 0,441 \text{ QALY}$.

Abbildung 13: Veranschaulichung des Markov-Modells für das Fallbeispiel Prävention mit Angabe der knotenspezifischen Parameter. Software: TreeAge Professional for Health Care (Fa. TreeAge, Williamstown, MA, USA).



p... = (Bedingte) Wahrscheinlichkeit. u... = Nutzwert. # = Wahrscheinlichkeitsverteilung wird an dieser Stelle automatisch zu 1 ergänzt.

Abbildung 14: Analysiertes Markov-Modell für das Fallbeispiel Prävention mit Angabe der erwarteten qualitätskorrigierten Lebensjahre (QALY) je Handlungsalternative für einen Zeitraum von 20 Jahreszyklen. Software: TreeAge Professional for Health Care (Fa. TreeAge, Williamstown, MA, USA).



FP = Falschpositiv. P = (Bedingte) Wahrscheinlichkeit. U = Nutzwert.

3.4.2.2 Sensitivitätsanalysen

In der EA versteht man unter einer Basisfallanalyse die Analyse eines Entscheidungsbaums oder eines Markov-Modells unter Verwendung bestimmter, nicht variiertes Ausprägungen für die Modellparameter. Eine Variation dieser Parameter ist im Rahmen von Sensitivitätsanalysen möglich, deren Ergebnisse dann mehr oder weniger vom Ergebnis der Basisfallanalyse abweichen. Sensitivitätsanalysen sind mathematische Verfahren, die die Auswirkung von Modellannahmen und deren Veränderungen auf den Entscheidungsausgang untersuchen. Variiert werden können dabei:

- Strukturelle Annahmen.
- Ereigniswahrscheinlichkeiten.
- Bewertungen der Gesundheitszustände.

Unter strukturelle Annahmen fällt die Struktur des Entscheidungsbaums bzw. des Markov-

Modells. Welche Verzweigungen in einem Entscheidungsbaum existieren oder nicht und welche Übergangswege zwischen den Gesundheitszuständen in einem Markov-Modell zugelassen sind und welche nicht, kann entscheidend den Ausgang der Analyse bestimmen. So enthalten beispielsweise entscheidungsanalytische Modelle zu chronischen Erkrankungen oft keinen Übergangsweg vom kranken in den gesunden Zustand. Hier wäre zu prüfen, inwieweit das Ergebnis der EA von einer zugelassenen Heilung abhängt. Ein ähnliches Problem besteht bei der Übernahme von Daten aus randomisierten klinischen Verlaufsstudien: In einer Studie von Rogers et al.⁵⁷ ergibt sich eine nicht unwesentliche Zahl von Alzheimer-Patienten, die im Verlauf der Therapie vom Zustand „moderate Symptomatik“ in den Zustand „milde Symptomatik“ wechseln, obwohl eher von einer progressionsverzögernden Wirkung der verabreichten Medikamente ausgegangen wird, als von einer Wirkung, die den Zustand akut verbessert⁵⁷. Da die Messfehlerrate bei den verwendeten kognitiven Messinstrumenten relativ hoch ist, kann nicht ausgeschlossen werden, dass die angesprochenen Phänomene zum Teil Artefakte darstellen.

Die Variation von Ereigniswahrscheinlichkeiten in Entscheidungsbäumen oder von Übergangswahrscheinlichkeiten in Markov-Modellen ist die häufigste Anwendung von Sensitivitätsanalysen. Da diese Parameter oft aus klinischen oder epidemiologischen Studien statistisch geschätzt werden, sind sie mit einem gewissen Grad an Unsicherheit behaftet. So werden Sensitivitätsanalysen oft über die Spannweite von angenommenen minimalen und maximalen Werten durchgeführt oder beispielsweise über den Bereich von 95 %-Konfidenzintervallen. Ein besonders wichtiger Fall einer Sensitivitätsanalyse von Wahrscheinlichkeiten ist die Variation der Krankheitsprävalenz. Sie wird weiter unten an einem Beispiel veranschaulicht.

Auch die Bewertung der Gesundheitszustände kann einer Sensitivitätsanalyse unterzogen werden. So können etwa in einer Kosten-Nutzwert-Analyse die QALY-Werte für bestimmte Gesundheitszustände variiert werden. In Kosten-Wirksamkeits-Analysen bietet sich die Variation der Preisstruktur an, um den Effekt zukünftiger Preisentwicklungen auf das Ergebnis der EA abschätzen zu können.

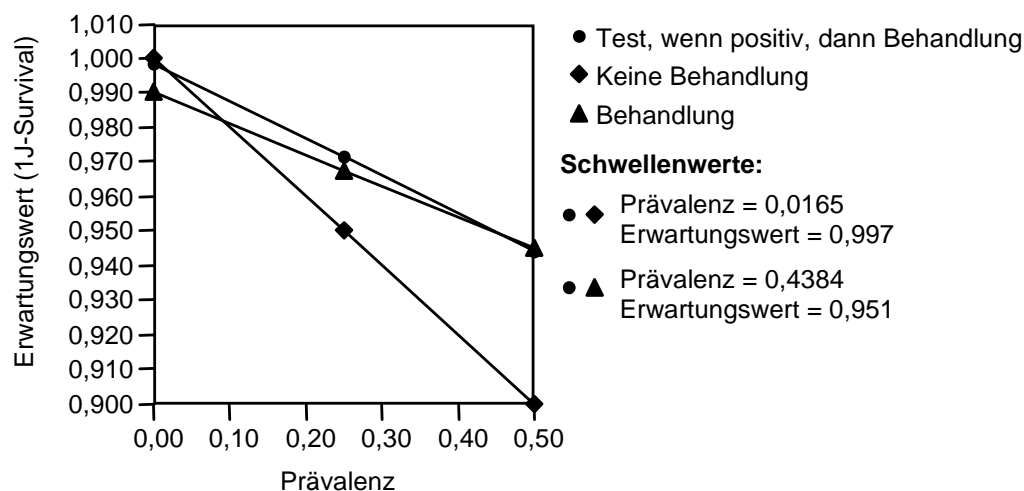
Ziel der Sensitivitätsanalyse ist neben der Prüfung der Stabilität der Entscheidung die Identifizierung von Parametern, die einen starken Einfluss auf das Ergebnis der EA haben. Werden solche Parameter identifiziert, so kann eine intensivere Forschung bezüglich dieser Parameter zu einer sichereren Einschätzung des Ergebnisses verhelfen. Ein besonderer Typ der Sensitivitätsanalyse ist die Schwellenwertanalyse: Unsichere Einflussparameter werden über einen bestimmten Bereich variiert, um den Wert dieser Parameter zu bestimmen, für den sich die Entscheidung ändern würde. Die Ermittlung von Schwellenwerten ist von besonderer Bedeutung für die Kliniker, um unter gegebenen Umständen schnell entscheiden zu können, welche Handlungsalternative unter gegebenen Umständen die optimale ist. Eine weitere Sonderform ist die Extremwertanalyse. Hier werden alle unsicheren Einflussparameter so gewählt, dass sie sich maximal für bzw. gegen eine bestimmte Entscheidung auswirken. Damit kann geprüft werden, ob eine Entscheidung auch unter extremen Annahmen stabil bleibt. Das Problem der Extremwertanalysen liegt in der Tatsache, dass bei vielen der nichttrivialen Entscheidungsprobleme eine Extremwertanalyse lediglich zu dem Ergebnis führt, dass die optimale Entscheidung von dem Szenario abhängt, da die Annahmen sehr konservativ zusammengeführt werden. Ferner kann es unter Umständen Schwierigkeiten bereiten, systematisch zu identifizieren, welche Parameterkonstellationen für oder gegen bestimmte Entscheidungen sprechen.

Je nachdem, ob ein oder mehrere Parameter simultan variiert werden, spricht man von Einweg- oder Mehrweg-Sensitivitätsanalysen. Eine weitere Einteilung der Sensitivitätsanalysen erfolgt

nach deterministischen versus probabilistischen Sensitivitätsanalysen. In deterministischen Sensitivitätsanalysen werden einzelne oder mehrere unsichere Einflussparameter jeweils über einen vorgegebenen Bereich variiert und die Ergebnisparameter in Abhängigkeit der Einflussparameter dargestellt. Die probabilistische Sensitivitätsanalyse wird auch verteilungsorientierte Sensitivitätsanalyse genannt und ist ein entscheidungsanalytisches Simulationsverfahren, dem folgendes Prinzip zugrunde liegt: Den einzelnen Parametern des entscheidungsanalytischen Modells werden Verteilungen statt feste Werte zugeordnet. Unter Verwendung eines Zufallsgenerators können dann für jeden unsicheren Einflussparameter Werte zufällig und wiederholt aus den zugehörigen Verteilungen gezogen werden. Im Rahmen einer Monte-Carlo-Simulation „laufen“ dann verschiedene hypothetische Patienten oder Kohorten mit den ihnen zugeordneten Parameterwerten durch das Modell. Dies ermöglicht die gleichzeitige Berücksichtigung der Unsicherheit in allen Parametern. Es resultiert eine Wahrscheinlichkeitsverteilung der erwarteten Ergebnisparameter. Zu weiteren methodischen Einzelheiten sei auf die Literatur verwiesen^{11, 22}. Die Vorteile der probabilistischen Sensitivitätsanalyse liegen u. a. in der Möglichkeit, multiple Sensitivitätsanalysen auch über eine große Zahl von Parametern simultan durchzuführen und außerdem die Unsicherheit des Ergebnisses in Form von Streudiagrammen, Konfidenzbereichen oder so genannten Cost-Effectiveness Acceptability Curves (CEAC) darstellen zu können.

Hier erfolgt eine Darstellung der Möglichkeiten anhand des bereits verwendeten Beispiels. Zunächst soll exemplarisch eine Einweg-Sensitivitätsanalyse dargestellt werden. Angenommen, es besteht Unsicherheit gegenüber der Prävalenz der Erkrankung. Daher wird diese zwischen 0 und 0,5 variiert. Abbildung 15 zeigt das Diagramm der Sensitivitätsanalyse.

Abbildung 15: Einweg-Sensitivitätsanalyse für die Prävalenz im Beispiel.



1J-Survival = Ein-Jahres-Überlebenschance.

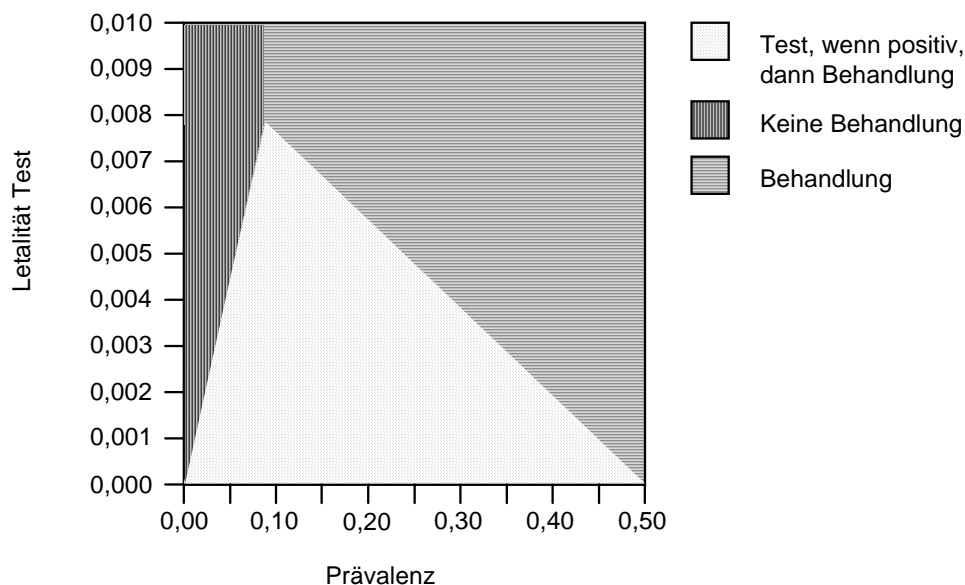
Die drei Graphen stellen die Erwartungswerte der drei verschiedenen Handlungsalternativen „Test“, „Keine Behandlung“ und „Behandlung“ als Funktion der Prävalenz dar. Auf der Rechtsachse ist der Parameter abgetragen, der variiert wird, in diesem Fall die Prävalenz. Auf der Hochachse lässt sich der Erwartungswert, also das erwartete 1J-Survival, in Abhängigkeit von der Prävalenz für jede Alternative ablesen. Zunächst ist festzustellen, dass bei allen Handlungsalternativen das Überleben bei zunehmender Prävalenz abnimmt. Dies ist plausibel: da keine Alternative vollständige Heilung bietet, sterben umso mehr Patienten, je mehr die

Krankheit haben. Am meisten von der Prävalenz beeinflusst ist das Schicksal der Patienten, wenn nicht behandelt wird (steil abfallender Graph). Demgegenüber beobachten wir den flachsten Graphen bei der Behandlung. Diese wirkt sozusagen nivellierend auf den Prävalenzeffekt. Die „Test“-Strategie liegt dazwischen. Der Schnittpunkt der Graphen „Behandlung“ und „Keine Behandlung“ repräsentiert der Schwellenwert für die Prävalenz bezüglich der Behandlung, falls kein Test zur Verfügung steht. Dieser Wert beträgt 0,0909 (im Diagramm nicht als Wert eingetragen). Das Schaubild zeigt, dass in weiten Teilen die „Test“-Strategie das höchste Überleben bringt. Dies war auch das Ergebnis der Basisfallanalyse, in dem sich bei einer Prävalenz von 0,02 für die „Test“-Strategie ein Erwartungswert von 0,9963 errechnet, während „Keine Behandlung“ zu einem Wert von 0,9960 und „Behandlung“ zu einem Wert von 0,9882 führen. Allerdings zeigt sich auch, dass es unter einem Schwellenwert der Prävalenz von 0,0165 besser ist, nicht zu behandeln als zu testen. Der Grund dafür liegt einmal darin, dass es bei einer sehr geringen Krankheitswahrscheinlichkeit keinen Sinn mehr macht, das Letalitätsrisiko des diagnostischen Tests auf sich zu nehmen. Aber auch wenn der diagnostische Test kein Risiko birgt, gibt es einen unteren Schwellenwert für die Überlegenheit der Strategie mit dem diagnostischen Test. Dieser liegt allerdings tiefer als bei gegebenem Letalitätsrisiko. In diesem Fall beträgt der Schwellenwert 0,0055; er ergibt sich durch Nullsetzen der Letalität in der EA. Der Grund dafür, warum Testen „schädlich“ sein kann, liegt darin, dass bei einer sehr geringen Prätestwahrscheinlichkeit (Prävalenz) selbst nach einem positiven nicht perfekten Test die Posttestwahrscheinlichkeit so gering ist, dass man nicht behandeln sollte. Es sei hier angemerkt, dass nicht der Test selbst den medizinischen Schaden mit sich bringt, sondern die Tatsache, dass die „Test“-Strategie so formuliert ist, dass nach einem positiven Test behandelt werden muss. In anderen Worten sollte man bei sehr kleinen Prävalenzen auch nach einem positivem Testergebnis nicht behandeln, was bedeutet, dass der Einsatz des Tests und die gewonnene Information sinnlos sind. Nur wenn der Test perfekt ist, also eine Sensitivität und eine Spezifität von 100 % hat, und der Test kein Risiko besitzt, dann erbringt die „Test“-Strategie immer den höchsten Erwartungswert. Dies ist plausibel, da (mit Ausnahme einer Prävalenz von 0 oder 1) die Unsicherheit bezüglich der Erkrankung durch die Testinformation ausgeräumt wird und für jeden einzelnen Patienten die richtige Entscheidung bezüglich der Behandlung getroffen werden kann. Der obere Schwellenwert beträgt 0,4384, d.h. über dieser Prävalenz sollen alle Patienten behandelt werden. Analog zum unteren Schwellenwert ist der Grund hierfür die Tatsache, dass bei einer hohen Prävalenz nicht riskiert werden soll, dass Patienten mit einem falsch negativen Testergebnis unbehandelt bleiben oder bereits durch den Test selbst sterben. Die Lage der Schwellenwerte hängt neben der Testletalität von den Behandlungskonsequenzen bei Kranken und bei Gesunden ab.

Variiert man nun zusätzlich den Parameter Testletalität, so lassen sich bereits aus Abbildung 15 einige Schlussfolgerungen ziehen. Bei geringerer Letalität würde sich der Graph der „Test“-Alternative nach oben verschieben, was zu einer Verbreiterung des Prävalenzbereichs führte, in dem sich der diagnostische Test als die überlegene Strategie erwiese. Umgekehrt führte eine Erhöhung der Testletalität zu einer Abwärtsverschiebung des „Test“-Graphs und der Bereich der Überlegenheit von „Test“ würde kleiner werden bis schließlich bei einem bestimmten Wert der Testletalität unabhängig von der Prävalenz ein diagnostischer Routinetest zu gefährlich wäre und die Strategien „Behandlung“ oder „Keine Behandlung“ (je nach Prävalenz) überlegen wären.

Um den Einfluss der Prävalenz der Erkrankung und der Letalität des diagnostischen Tests auf die Entscheidung graphisch darzustellen, bedient man sich der Zweiweg-Sensitivitätsanalyse (Abbildung 16).

Abbildung 16: Zweiweg-Sensitivitätsanalyse für Letalität und Prävalenz im Beispiel.



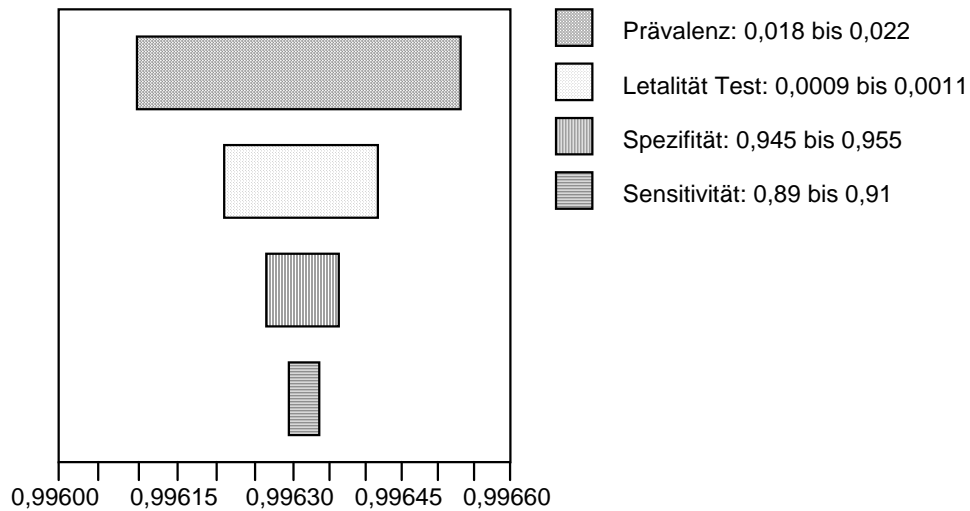
Dieser Graph zeigt keine Erwartungswerte mehr an, kennzeichnet jedoch für jede Kombination von Prävalenz und Letalität die optimale Handlungsalternative. In jedem der drei schraffierten Bereiche in Abbildung 16 ist jeweils eine andere Handlungsalternative den verbleibenden Handlungsalternativen überlegen. Die Grenzlinien beschreiben die Indifferenzkurven zwischen zwei Alternativen. Der Schnittpunkt aller drei Kurven repräsentiert die Konstellation der Parameter Prävalenz und Letalität, bei der alle drei Handlungsstrategien denselben Erwartungswert besitzen. Es erweist sich z.B. für eine Prävalenz von 0,05 und einer Letalität von 0,001 die „Test“-Strategie als überlegen, während bei einer Prävalenz von 0,01 und einer Letalität von 0,002 die Alternative „Keine Behandlung“ vorzuziehen wäre. Die Tatsache, dass bei zunehmender Letalität das Prävalenzintervall für den sinnvollen Testeinsatz immer kleiner wird, lässt sich anschaulich aus dem Diagramm ablesen: Die „Test“-Fläche läuft nach oben spitz zu. Bei einer Testletalität von über ca. 0,008 ist es besser, das Testen zu unterlassen, in diesem Bereich kann wieder der Prävalenzschwellenwert 0,09 bezüglich „Behandlung und „Keine Behandlung“ abgelesen werden (als Rechtswert der Grenzlinie zwischen den Alternativen „Behandlung“ und „Keine Behandlung“).

Eine beliebte graphische Veranschaulichung der Ergebnisse mehrerer Sensitivitätsanalysen ist das so genannte Tornadodiagramm, das in Abbildung 17 wiedergegeben ist. Es veranschaulicht eine Serie von Einweg-Sensitivitätsanalysen mehrerer Parameter und stellt sie gemeinsam in einem Diagramm dar. Horizontale Balken geben den Einfluss der Variation eines jeden untersuchten Parameters auf den Erwartungswert der jeweils besten Handlungsalternative an. Abbildung 17 zeigt für das Testbeispiel den Einfluss der Variablen Prävalenz, Letalität des diagnostischen Tests, Spezifität und Sensitivität auf das erwartete 1J-Survival. In der Legende sind die Variationsbereiche für die verschiedenen Parameter angegeben. Auf der horizontalen Achse sind die Erwartungswerte für das 1-Jahres-Survival angegeben.

Im Tornadodiagramm (Abbildung 17) zeigt sich, dass unter den untersuchten Parametern die Variation der Prävalenz den größten Einfluss und die Variation der Testsensitivität den geringsten Einfluss auf den Erwartungswert hat. Das 1J-Survival variiert zum Beispiel zwischen 0,99615 für den unteren Variationsbereich und 0,99660 für den oberen Variationsbereich der

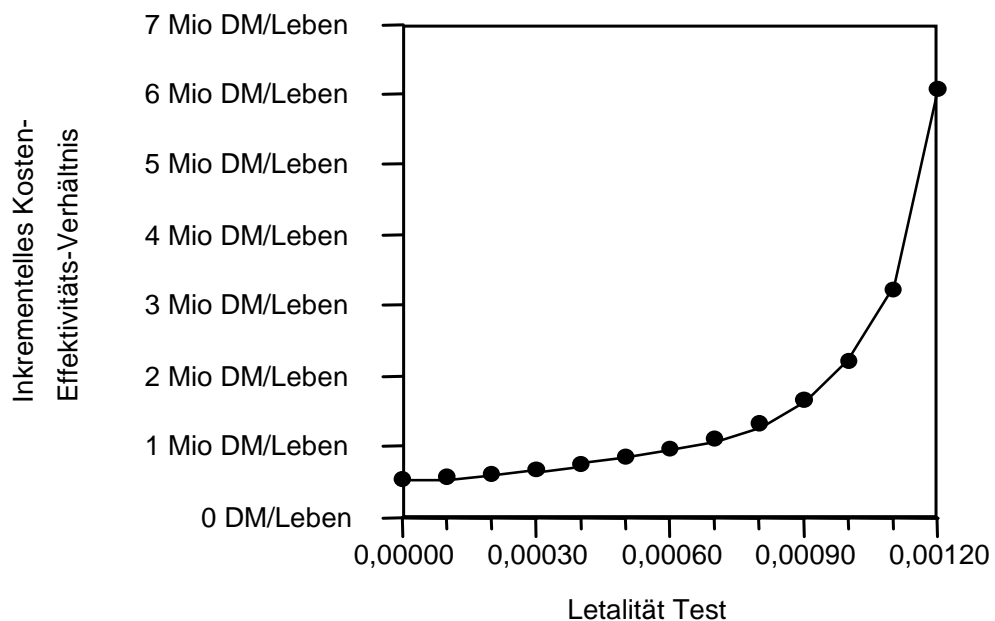
Prävalenz.

Abbildung 17: Tornadodiagramm für den Einfluss der Variablen Prävalenz, Letalität des diagnostischen Tests, Spezifität und Sensitivität auf die erwartete Ein-Jahres-Überlebenswahrscheinlichkeit im Beispiel.



Abschließend sei noch eine Sensitivitätsanalyse zum Kosteneffektivitätsverhältnis dargestellt: Abbildung 18 zeigt den Verlauf des Kosteneffektivitätsverhältnisses für den Vergleich der Alternativen „Test“ versus „Keine Behandlung“ in Abhängigkeit von der Testletalität. Bei zunehmender Letalität nimmt die medizinische Effektivität durch das diagnostische Testen immer mehr ab, so dass bei etwa gleich bleibenden Kosten das Kosteneffektivitätsverhältnis immer weiter ansteigt um für eine Effektivität von Null den Wert unendlich zu erreichen. Berücksichtigt man die Lebenserwartung des Patienten, so können die dargestellten Werte in die üblichere Einheit Kosten pro gewonnenes Lebensjahr umgerechnet werden, so wie es im obigen Beispiel erfolgt.

Abbildung 18: Sensitivitätsanalyse des Kosteneffektivitätsverhältnisses für die Testletalität im Beispiel.



3.4.2.6 Anwendungsbeispiel eines Markov-Modells

Ziel eines HTA-Berichts im Auftrag des DIMDI und des Bundesministeriums für Gesundheit und Soziale Sicherung (BMGS) ist die Evaluation der Wertigkeit des Einsatzes der genotypischen und phänotypischen HIV-Resistenzbestimmung im Rahmen der Behandlung von HIV-infizierten Patienten^{17, 18}. Um die Kosteneffektivität des Einsatzes genotypischer Resistenztests zu analysieren, wird von den Autoren des genannten HTA-Berichts ein entscheidungsanalytisches Markov-Modell für den deutschen Gesundheitskontext entwickelt, das in drei separaten Markov-Modellen (State Transition Models) den Krankheitsverlauf und die Behandlungskosten einer Kohorte HIV-Infizierter Personen unter drei verschiedenen Strategien simuliert. Diese verglichenen Strategien sind:

- *No Test*: Behandlung der HIV-Infektion ohne genotypische Resistenztests und Umstellung der Therapie nach konventionellen Kriterien (z.B. Therapieanamnese).
- *Test After First Failure*: Resistenztests vor jeder Therapieumstellung, d.h. erstmals nach Versagen der Initialtherapie.
- *Test Always*: Genotypisierung vor jeder Therapiezusammenstellung, d.h. auch schon vor der Initialtherapie.

Die Markov-Kohorte startet im Zustand der Therapienaivität und beginnt die erste antiretrovirale Therapiekombination (ART1). Bei initialem Therapieerfolg (Suppression der Viruslast unter die Nachweisgrenze von 500 Kopien / ml nach sechs Monaten) wird ART1 fortgesetzt. Bei initialem Therapieversagen (Nichtunterschreiten der Nachweisgrenze bis Monat 6) erfolgt die Umstellung auf die zweite Kombinationstherapie (ART2). Bei initialem Erfolg aber späterem Versagen von ART1 (Wiederanstieg der Viruslast über die Nachweisgrenze) erfolgt ebenfalls die Umstellung auf ART2. Dieser Prozess setzt sich fort bis zum Versagen der vierten Kombination (ART4), nach der die therapeutischen Optionen erschöpft sind, basierend auf der Zahl der zurzeit verfügbaren Medikamente und ihrer durch Unverträglichkeiten und Kreuzresistenzen begrenzten Kombinationsmöglichkeiten. Zu jedem Zeitpunkt sind zudem der Übergang in das AIDS-Stadium sowie der Tod durch andere Ursachen möglich. Vom AIDS-Stadium ist lediglich ein Übergang zu Tod durch AIDS möglich. Als Zykluslänge wird ein Zeitraum von sechs Monaten definiert, da dies die Periode ist, nach der initialer Erfolg oder Versagen einer Kombinationstherapie beurteilt werden können. Die Übergangswahrscheinlichkeiten werden aus publizierten klinischen Studien und Beobachtungsstudien entnommen. Kostendaten entstammen einer deutschen Studie und deutschen Medikamentenpreisen. Es wird eine Kosteneffektivitätsanalyse aus der Perspektive des gesetzlichen Krankenversicherungssystems durchgeführt. Outcomes sind Restlebenserwartung, Lebenszeitkosten und das diskontierte inkrementelle Kosteneffektivitätsverhältnis (IKEV). Es werden deterministische und probabilistische Sensitivitätsanalysen (Monte-Carlo-Simulation) durchgeführt.

In der Basisfallanalyse kann gezeigt werden, dass genotypische Resistenztestung nach einem Therapieversagen die Lebenserwartung um neun Monate erhöht und 16.406 Euros kostet. Nach Diskontierung mit jährlich 3 % wird ein IKEV von 22.510 Euro / LJ (LJ = Lebensjahr) ermittelt. In den Extremwertanalysen variiert das IKEV zwischen 16.512 Euro / LJ und 42.900 Euro / LJ. In der Grenzwertanalyse wird berechnet, dass der Einsatz genotypischer Resistenztestung vor initialer Therapie kosteneffektiv sei, wenn die Wahrscheinlichkeit eines Therapieversagens um mindestens 36 % gesenkt würde.

Dieses Beispiel zeigt den Einsatz eines entscheidungsanalytischen Markov-Modells zur Evaluation des Einsatzes eines diagnostischen Testverfahrens zur Optimierung einer Therapie

im nationalen Kontext. Sowohl deterministische als auch probabilistische Sensitivitätsanalysen, Extremwert- und Schwellenwertanalysen werden hier durchgeführt, um die Robustheit der Analyseergebnisse zu evaluieren und bestimmte Fragestellungen zu beantworten. Details dieser Studie können im publizierten Fachartikel¹⁷ und dem HTA-Bericht¹⁸ nachgelesen werden.

3.5 Ergebnisse

(Verfasser: U. Siebert, B. Bornschein)

3.5.1 Systematische Literaturrecherche

Die Literatursuche dient der Identifikation von Arbeiten, die einen methodischen Beitrag zur Übertragbarkeit von Evidenz unter Anwendung entscheidungsanalytischer Verfahren leisten oder von Beispielen für die Anwendung einer solchen Methodik berichten.

Im Folgenden sind die einzelnen Rechenschritte sowie die jeweilige Anzahl der Treffer zur Identifikation relevanter Literatur zu Problemen und zu Lösungsansätzen im Rahmen entscheidungsanalytischer Evidenzübertragung in den Datenbanken MEDLINE (Tabelle 6) und EconLit (Tabelle 7) zusammengestellt. Tabelle 8 zeigt die Ergebnisse der weiter gefassten Literaturrecherche zum spezifisch deutschen Kontext.

Tabelle 6: Rechenschritte und –ergebnisse in der Literaturdatenbank MEDLINE.

Rechenschritt	Suchbegriff	Treffer
#1	decision analy\$.af	2.088
#2	decision tree.af	812
#3	decision supp\$.af	6.124
#4	markov\$.af NOT markov\$.au	2.583
#5	decision mak\$.af NOT decision mak\$.jn	53.671
#6	1 OR 2 OR 3 OR 4 OR 5	61.616
#7	clinical.af	1.101.755

Fortsetzung Tabelle 6

#8	epidemiolog\$.af	158.400
#9	(health technology OR hta).af	932
#10	(cost\$ AND effect\$).af	58.257
#11	7 OR 8 OR 9 OR 10	1.269.483
#12	11 AND (evaluat\$ OR assessment).af	292.850
#13	generali#ation.ti,ab	5.036
#14	\$national.ti	28.749
#15	countr\$.ti	10.759
#16	transfer.ti,ab	107.482
#17	13 OR 14 OR 15 OR 16	151.642
#18	6 AND 12 AND 17	158
#19	18 AND (ge\$ OR en\$ OR fr\$).lg	156
#20	limit 19 to yr=1985-2001	150
#21	20 NOT genetic\$.af	145

Das \$-Zeichen ist ein Trunkierungssymbol, d.h. es steht für beliebige (auch leere) Zeichenfolgen. Das #-Zeichen im Suchbegriff symbolisiert ein einzelnes beliebiges Zeichen. Die Ziffern des Suchbegriffs bezeichnen das Ergebnis des jeweiligen Rechenschritts. Die Abfrage erfolgt am 17.12.2001.

Tabelle 7: Rechenschritte und –ergebnisse in der Literaturdatenbank EconLit.

Rechenschritt	Suchbegriff	Treffer
#1	decision analy\$.af	194
#2	decision tree.af	26
#3	decision supp\$.af	322
#4	markov\$.af NOT markov\$.au	5.939
#5	decision mak\$.af NOT decision mak\$.jn	8.070
#6	1 OR 2 OR 3 OR 4 OR 5	14.340
#7	clinical.af	316
#8	epidemiolog\$.af	143
#9	(health technology OR hta).af	11
#10	(cost\$ AND effect\$).af	99.014
#11	7 OR 8 OR 9 OR 10	99.326
#12	11 AND (evaluat\$ OR assessment).af	6.804
#13	generali#ation.ti,ab	942
#14	\$national.ti	4.993
#15	countr\$.ti	11.927
#16	transfer.ti,ab	3.910
#17	13 OR 14 OR 15 OR 16	21.426
#18	6 AND 12 AND 17	8
#19	limit 18 to yr=1985-2001	8

Das \$-Zeichen ist ein Trunkierungssymbol, d.h. es steht für beliebige (auch leere) Zeichenfolgen. Das #-Zeichen im Suchbegriff symbolisiert ein einzelnes beliebiges Zeichen. Die Ziffern des Suchbegriffs bezeichnen das Ergebnis des jeweiligen Rechenschritts. Die Abfrage erfolgt am 18.12.2001.

Tabelle 8: Rechenschritte und –ergebnisse in der Literaturdatenbank MEDLINE für den deutschen Kontext.

Rechenschritt	Suchbegriff	Treffer
#1	decision anal\$.af	2.089
#2	german\$.af	1.034.458
#3	1 AND 2	63
#4	limit 3 to yr=1985-2001	57

Das \$-Zeichen ist ein Trunkierungssymbol, d.h. es steht für beliebige (auch leere) Zeichenfolgen. Die Ziffern des Suchbegriffs bezeichnen das Ergebnis des jeweiligen Rechenschritts. Die Abfrage erfolgt am 18.12.2001.

Das MEDLINE-Suchergebnis von 145 Zitaten wird auf doppelte Zitate untersucht, dabei werden 16 doppelte Datensätze identifiziert. Nach der Elimination der Redundanzen verbleiben 129 Literaturstellen. Die EconLit-Recherche ergibt acht Treffer, die einzeln auf ihre Relevanz

überprüft wurden. Die weiter gefasste Recherche für den spezifisch deutschen Kontext erbringt 57 Treffer.

Alle Treffer werden anhand ihrer Zusammenfassungen oder ihrer Volltexte auf ihre Relevanz für die Fragestellung untersucht. Keine der identifizierten Originalarbeiten entspricht den oben genannten Kriterien bezüglich relevanter Aspekte für die Fragestellung. Auch Hinweise aus Artikeln bzw. ihre Literaturlisten ergeben keine weiteren relevanten Dokumente.

Eine Durchsicht gesundheitsökonomischer Standardlehrbücher und Lehrbuchkapitel ergibt, dass das recherchierte Thema lediglich in den Lehrbüchern von Drummond et al. und Schöffski et al. Eingang gefunden hat^{24, 61}.

Das Lehrbuch von Drummond et al.²⁴ weist auf die Möglichkeit hin, Übertragung von Evidenz im Zusammenhang mit gesundheitsökonomischen Evaluationen über Entscheidungsbäume zu modellieren. Es erfolgen allgemeingültige Hinweise auf die Übertragung relevanter Parameter. Zur Illustration werden zwei Beispiele aus der Literatur aufgeführt, davon einer für die Verwendung der EA²³. Die Originalveröffentlichungen beider Beispiele sind über die Literaturabfrage nicht identifiziert worden.

Das Lehrbuch von Schöffski et al.⁶¹ widmet der Übertragbarkeit von Evidenz ein eigenes Kapitel. Es beruft sich dabei in weiten Teilen auf Welte und Leidl⁷⁷. Zu übertragende Parameter werden ausführlich diskutiert und auch entscheidungsanalytische Lösungsansätze genannt. Dabei wird festgestellt, dass das Thema der Übertragbarkeit sowohl empirisch als auch konzeptionell überraschend wenig bearbeitet werde. Als Anwendungsbeispiel wird wiederum die genannte Arbeit von Drummond et al.²³ angeführt. Dieses Beispiel wird außerdem in einer weiteren Publikation von Greiner et al.³² aufgegriffen, die neben den Problemen bei der Interpretation gesundheitsökonomischer Daten aus multinationalen Studien und der Übertragbarkeit solcher Ergebnisse auf den Kontext eines bestimmten Ziellands den Stand der derzeitigen Praxis sowie mögliche Lösungsansätze diskutiert. Als zentrale Bereiche werden die Kostenstrukturen, die epidemiologischen und demographischen Rahmenbedingungen (unterschiedlicher Casemix) und die gesundheitssystemspezifischen Anreizstrukturen (z.B. DRG-bezogene Vergütung: DRG = Diagnosis-related groups) sowie deren Bedeutung für die Ressourcenverbräuche genannt.

In dieser Arbeit werden außerdem verschiedene Lösungsansätze zur Übertragung gesundheitsökonomischer Evidenz zwischen verschiedenen Ländern diskutiert. Es wird ein Lösungsmodell in Form eines Flussdiagramms vorgestellt, das die verschiedenen methodischen Optionen zur Übertragung multinationaler Studiendaten zur Beantwortung nationaler Studienfragen aufzeigt. Die Autoren empfehlen den entscheidungsanalytischen Ansatz als die Methode der Wahl, sofern Studiendesign und Fragestellung dies erlauben. Der Schwerpunkt der Arbeit liegt auf der Übertragbarkeit der Evidenz zu ökonomischen Daten, insbesondere bei der Verwendung von Daten aus multinationalen Studien. Ansätze zur Anpassung klinischer Studienergebnisse werden nicht vorgestellt.

Vier weitere Artikel mit anderer Zielsetzung wiesen zumindest teilweise relevante Aspekte der vorgegebenen Fragestellung auf^{34, 51, 71, 74}.

Eine Arbeitsgruppe der World Health Organisation (WHO) berichtet von einer multizentrischen randomisierten klinischen Studie mit ökonomischem Appendix, die in Argentinien, Kuba, Saudi-Arabien und Thailand durchgeführt wird mit dem explizit formuliertem Ziel, zusätzlich eine Übertragbarkeit der Ergebnisse auf ausgewählte weitere Länder zu überprüfen. Dabei handelt es sich um Südafrika, Gambia, Zimbabwe, Indonesien und Bangladesch⁵¹. Über die beabsich-

tigte Methodik bei der Prüfung der Übertragbarkeit werden keine Angaben gemacht. Die Ergebnisse der Hauptstudie liegen inzwischen veröffentlicht vor⁷⁴, die Publikation der Zusatzstudie zur Übertragbarkeit der Ergebnisse steht jedoch offensichtlich noch aus.

In einem weiteren Artikel versucht eine französische Arbeitsgruppe anhand einer Literaturliteraturauswertung festzustellen, ob Evidenz aus gesundheitsökonomischen Evaluationsstudien zur adjuvanten Therapie bei Frauen mit Brustkrebs auf das französische Gesundheitssystem übertragen werden kann. Von 234 Treffern bei der Abfrage von Literaturdatenbanken entspricht keine einzige Studie den von den Forschern vorgegebenen Qualitätskriterien, die für eine Übertragbarkeit als Voraussetzung gelten sollten⁷¹.

In einer systematischen Arbeit von Haycox et al. wird für vier europäische Länder, die sich auf einen einheitlichen grundlegenden klinischen Behandlungsalgorithmus für Erkrankungen des oberen Gastrointestinaltrakts verständigt haben, eine an die EA angelehnte Beschreibung der jeweiligen nationalen Kosten erarbeitet. Sie werden mit der Situation in Großbritannien als Referenzsituation verglichen. Dabei werden nationale Besonderheiten der Gesundheitssysteme speziell bei der Preisbildung im Arzneimittelbereich, wie z.B. Zuzahlungsregelungen berücksichtigt³⁴. Diese Studie kann als Vorläufer eines entscheidungsanalytischen Übertragungsmodells betrachtet werden.

Zusammenfassend ist festzuhalten, dass entscheidungsanalytische Ansätze zur Übertragung von Evidenz aus ausländischen Studien auf den nationalen Kontext in der Literatur weitgehend fehlen. Das Grundkonzept für derartige Forschungsansätze kann in einzelnen Fällen bereits andeutungsweise erkannt werden. Im Gegensatz dazu werden entscheidungsanalytische Verfahren von modernen Standardlehrbüchern zunehmend propagiert.

3.5.2 Literaturbeispiele zu Standardsituationen notwendiger Evidenzübertragung

Die erste Forschungsfrage dieses HTA-Berichts richtet sich auf Standardsituationen, bei denen internationale Evidenz auf den Kontext eines bestimmten Gesundheitssystems zu übertragen ist. Zur Beantwortung dieser Frage werden im Folgenden Standardsituationen, bei denen internationale Evidenz auf den Kontext eines bestimmten Gesundheitssystems zu übertragen ist, anhand von Literaturbeispielen dargestellt.

Es existiert eine Vielzahl an denkbaren Konstellationen, die für Entscheidungsträger des Gesundheitssystems von Interesse im Hinblick auf eine Evidenzübertragung auf deutsche Verhältnisse sein dürften. Aus dieser Vielzahl von Möglichkeiten sollen im Folgenden einzelne Beispiele besonders häufig anzutreffender Situationen herausgegriffen und exemplarisch dargestellt werden.

3.5.2.1 Effektivität therapeutischer Verfahren

Von zentraler Bedeutung für den medizinischen Fortschritt sind die Entwicklung neuer Therapieprinzipien und ihre Evaluation bezüglich ihrer Wirksamkeit. Die entsprechenden Studien werden häufig als international angelegte randomisierte kontrollierte Multicenterstudien (RCT) durchgeführt. Wie weit sich die Ergebnisse auch auf die Bevölkerung von Ländern generalisieren lassen, die an den Studien nicht teilgenommen haben, kann nur unter Zusatzannahmen beantwortet werden. Insbesondere ist die Übertragung der Ergebnisse von RCT auf die realen Versorgungssituationen Gegenstand anhaltender Kritik²²¹. So werden die meisten Arbeiten zur Sekundärprävention koronarer Herzkrankheit mit Statinen zur Cholesterinsenkung an Männern durchgeführt, Empfehlungen werden aber auch für die Behandlung von Frauen abgeleitet⁴⁶.

In einer kürzlich publizierten Metaanalyse von Boyle et al.⁹ zur Effektivität von Permixon in der

Behandlung der benignen Prostatahyperplasie demonstrieren die Autoren den Wert des Einschlusses von großen multizentrischen Beobachtungsstudien neben den kontrollierten klinischen Studien, da diese zusätzliche Informationen liefern können. Im Allgemeinen herrscht derzeit unter den Experten noch Uneinigkeit darüber, ob Ergebnisse von Beobachtungsstudien mit großen Patientenzahlen gegenüber den Ergebnissen von kleinen RCT bevorzugt und zur klinischen Entscheidungsfindung herangezogen werden sollen⁶³. Große Beobachtungsstudien haben gegenüber den RCT den Vorteil, dass sie Information darüber geben können, in wieweit sich die beobachtete wissenschaftliche Evidenz zur Effektivität, gemessen unter den spezifischen Bedingungen in klinischen Studien auch bei einem Einsatz in der Routineversorgung widerspiegelt^{13, 40, 47}. Im Mittel wird eine geringere Studienqualität für Beobachtungsstudien im Vergleich zu RCT gefunden. Kunz und Oxman⁴⁰ berichten, dass Ergebnisse aus nicht-randomisierten Studien häufig den Behandlungseffekt aufgrund einer besseren „Baseline“-Prognose in der Gruppe der behandelten Patienten überschätzen. Andere Autoren schildern, dass die Erklärung der Heterogenität das wichtigste Ziel einer Metaanalyse sein kann^{7,30, 31, 35, 42, 55, 58, 59, 73}.

3.5.2.2 Effektivität diagnostischer Verfahren

Die Effektivität neuer diagnostischer Verfahren wird zunächst an ausgewählten Populationen mit hohen Prävalenzen für die relevante Krankheit geprüft. Bei der Verwendung eines derart evaluierten Tests in anderen Bevölkerungen, wie z.B. in der Allgemeinbevölkerung des gleichen Lands, in der eines Nachbarlands oder in besonderen Risikogruppen können die diagnostischen Effektivitätsdaten zum Teil erheblich variieren. Eine direkte Übertragung der Ergebnisse hinsichtlich der Effektivität des diagnostischen Tests ist schwierig. Die Zusammensetzung der Studienpopulation hinsichtlich der Krankheitsprävalenz stellt deshalb einen wichtigen Parameter für die Übertragbarkeit von Ergebnissen dar. Das betrifft zum Beispiel primär prävalenzabhängige Outcomeparameter wie positive Vorhersagewahrscheinlichkeiten (PPV) und negative Vorhersagewahrscheinlichkeiten (NPV). Mit steigender Prävalenz steigen jeweils auch die ermittelten PPV. Das bedeutet, dass an Hochrisikostudienpopulationen ermittelte PPV nicht auf Screeningpopulationen (Normalbevölkerung) übertragbar sind. So berichten beispielsweise Schiffman et al. über die Ergebnisse für einen Test auf humanes Papillomavirus (HPV) von 88 % Sensitivität und 89 % Spezifität bei Frauen eines Endemiegebiets für Zervixkarzinome in Costa Rica⁶⁰. Inwieweit diese Ergebnisse bei einer Anwendung in der deutschen weiblichen Allgemeinbevölkerung mit möglicherweise ganz anderen Prävalenzen für HPV und Zervixkarzinomen reproduziert werden können, kann aus diesen Daten nicht abgeleitet werden.

Auch die nicht prävalenzabhängigen Testgütekriterien Sensitivität und Spezifität werden durch ergebnisverzerrende systematische Fehler beeinflusst. McCrory et al.⁴⁸ können in ihrer Metaanalyse zur Effektivität des Papanicolaou-Tests eine deutliche Ergebnisverzerrung von Sensitivität und Spezifität durch einen Work-Up-Bias zeigen. Dieser systematische Fehler entsteht, wenn die Anwendung des Referenzstandards in Abhängigkeit von den Ergebnissen des Screeningtests erfolgt. Es wird gezeigt, dass bei vorliegendem Work-Up-Bias die Sensitivität eher überschätzt und die Spezifität eher unterschätzt wird. Da der Work-Up-Bias mit steigender Prävalenz zunimmt, wird die Ergebnisverzerrung in hochprävalenten (Hochrisiko-)Populationen stärker ausfallen, als in Screeningpopulationen niedriger Prävalenz⁴⁸.

Weitere Faktoren spielen eine entscheidende Rolle wie zum Beispiel die Wahl des zytologischen Schwellenwerts für einen positiven Befund, die Durchführung eines Vergleichs mit einem anerkannten Goldstandard und falls nicht bei allen Testobjekten eine Verifikation durch

einen Goldstandard stattfindet, die unverzerrte Auswahl der Testobjekte, die mit einem Goldstandard verifiziert werden.

Ein eindrucksvolles Beispiel ist auch hier die Ergebnispräsentation der Sensitivität und der Spezifität des Papanicolaou-Tests im Zervixkarzinomscreening. McCrory et al.⁴⁸ berichten von einer gepoolten Sensitivität von 51 % für den Papanicolaou-Test bei dem Schwellenwert ASCUS (Atypical Squamous Cells of Undetermined Significance). In ihrer Metaanalyse werden sowohl Studien gepoolt, die eine Hochrisikopopulation untersuchen als auch Arbeiten mit Normalrisikopopulationen. Es werden Studien ausschließlich mit Goldstandardverifikation berücksichtigt, jedoch nicht ein Verifikationsbias ausgeschlossen. In einer weiteren Publikation derselben Forschergruppe⁵² werden basierend auf dem Studienpool der Metaanalyse von McCrory et al.⁴⁸ Sensitivitäten und Spezifitäten des Papanicolaou-Tests disaggregiert für (1) Populationen mit geringem / hohem Erkrankungsrisiko, (2) verschiedene Testschwellenwerte und (3) um einen Verifikationsbias bereinigt dargestellt. Es zeigt sich eine erhebliche Variation von Sensitivität und von Spezifität des Tests unter Berücksichtigung der genannten Faktoren. Im Vergleich fällt die Sensitivität des Papanicolaou-Tests mit 47 (Testschwellenwert: LSIL (Low-Grade Squamous Intraepithelial Lesion) versus 69 % deutlich geringer aus unter Berücksichtigung des Verifikationsbias und unter ausschließlicher Verwendung von Studien mit Normalrisikopopulationen.

3.5.2.3 Prognose und Risikoabschätzung

Die Identifikation von Subpopulationen mit besonders hohem Risiko in der Absicht, diese entweder von bestimmten medizinischen Prozeduren auszuschließen, ist wesentlicher Teil ärztlicher Tätigkeit. Diese Risikoabschätzung findet explizit z.B. bei Untersuchungen zur Operationsvorbereitung oder implizit bei der Verordnung von Arzneimitteln bei der Berücksichtigung von Kontraindikationen statt. Unter diesen Umständen wird von einer Allgemeingültigkeit der Evidenz ausgegangen. Um diese Risikobewertung auf Patientencharakteristika basieren zu können, werden so genannte prognostische Scores erstellt. Beispielsweise wird in einer Prognosestudie zur Abschätzung des Operationsrisikos einer koronaren Bypassoperation das Risiko als Funktion der Patienteneigenschaften bei für die Operation vorgesehenen Patienten bestimmt⁵⁴. Die Autoren erwägen, das berechnete Prognoseergebnis als „Electronic Second Opinion“ anderen Ärzten verfügbar zu machen. Sie berichten jedoch, dass das in ihrem herzchirurgischen Spezialzentrum erarbeitete Prognosemodell möglicherweise nicht ohne weiteres auf andere, weniger spezialisierte Zentren übertragen werden kann bzw. dass eine solche Übertragung von Evidenz nur mit größter Vorsicht erfolgen dürfe.

3.5.2.4 Kosteneffektivität medizinischer Verfahren

Das wohl derzeit größte Interesse bei der Evidenzübertragung scheint sich auf den Bereich der Kosteneffektivität medizinischer Technologien zu richten. Neben der Übertragung von Daten zur medizinischen Effektivität sind auch Kostendaten zu übertragen⁸⁴. Gerade diese sind in hohem Maß abhängig vom Versorgungskontext und dem jeweiligen Gesundheitssystem, so dass hier besondere Probleme bestehen. So ergeben sich bei einem Vergleich des Kosteneffektivitätsverhältnisses der Behandlung von aneurysmatischen Subarachnoidalblutungen mit dem Medikament Tirilazad-Mesylat in fünf europäischen Ländern Werte zwischen 11.450 und 244.133 US\$ pro verhindertem Todesfall⁷⁸.

Bei der Bewertung der Kosteneffektivität diagnostischer Verfahren ist die Prävalenz der nachzuweisenden Krankheit meist ein kritischer Parameter. In einer Studie an 26 französischen Bluttransfusionszentren wird untersucht, ob das neu eingeführte Screening auf Antikörper

gegen Hepatitis C gegenüber der bisherigen Strategie (nur Bestimmung von Antikörpern gegen Hepatitis B und Transaminasen) kosteneffektiv ist. Es zeigt sich, dass die neu eingeführte Screeningstrategie zwar kosteneffektiv erscheint, jedoch ihr medizinischer und ihr ökonomischer Wert sehr stark von der Prävalenz abhängen. Die Autoren kommen zu dem Schluss, dass Transplantationszentren in Regionen Frankreichs in Abhängigkeit von den Prävalenzen für Hepatitis C gegebenenfalls ihre bisherige Screeningstrategie ändern sollten¹⁶.

3.5.2.5 Wirksamkeit versus Effektivität unter Routinebedingungen

Zentraler Kritikpunkt an der Evidenz-basierten Medizin ist die Frage nach einer Übertragbarkeit der im Rahmen von RCT an größeren Gruppen gewonnenen Erkenntnisse auf die einzelnen Patienten des klinischen Behandlungsalltags. Die Kritik ist berechtigt insbesondere vor dem Hintergrund umfangreicher Ein- und Ausschlusskriterien bei der Durchführung der RCT, sowie möglicherweise anderer Risikokonstellationen des individuellen Patienten. So stellt sich beispielsweise die Frage, ob bei einem Patienten mit vorgeschädigter Leber die evidenzgesicherte Behandlung mit Vitamin-K-Antagonisten ebenfalls vorteilhaft ist^{21, 46}.

Das artifizielle Design von RCT, das zur Optimierung des Wirksamkeitsnachweises dient, führt zu anderen quantitativen Ergebnissen der Effektivität, als sie unter Realbedingungen im Alltag erreichbar sind, so dass auch hier eine Übertragung bzw. Adjustierung der Evidenz notwendig sein wird².

3.5.2.6 Leitlinien

In den methodischen Standards der Entwicklung evidenzbasierter Leitlinien in Deutschland kommt der Sichtung der internationalen Literatur zur evidenzbasierten Medizin ein hoher Stellenwert zu⁸⁵. Allerdings sollen Leitlinien letztlich als klinische Handlungsempfehlungen Bezüge zum nationalen Kontext aufweisen bzw. für diesen entwickelt werden.

Eine Systematik zur Übertragung oder Anpassung von Leitlinien oder deren Komponenten von einem Kontext zum anderen erscheint grundsätzlich möglich, findet aber bislang wenig Eingang bei behandelnden Ärzten²⁹. Ein strukturierter und formaler Algorithmus zum Transfer der den Leitlinien zugrunde liegenden Evidenz würden die Transparenz und die Qualität der Leitlinien erhöhen sowie damit deren Akzeptanz und Verbreitung fördern.

Nach den von der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und der Ärztlichen Zentralstelle Qualitätssicherung (ÄZQ) gegebenen Definition sind Leitlinien systematisch entwickelte Feststellungen, um die Entscheidungen von Klinikern und Patienten über angemessene Gesundheitsversorgung für spezifische klinische Umstände zu unterstützen⁸⁶.

In ihren methodischen Empfehlungen „Leitlinie für Leitlinien“⁸⁷ klassifiziert die AWMF die Leitlinien für Diagnostik und Therapie in drei Kategorien. In der Entwicklungsstufe 1 einer Leitlinie (S1-Leitlinie) erarbeitet eine repräsentativ zusammengesetzte Expertengruppe der AWMF im informellen Konsens eine Leitlinie, die vom Vorstand der Fachgesellschaft verabschiedet wird. In der Entwicklungsstufe 2 (S2-Leitlinie) werden vorhandene S1-Leitlinien in einem der bewährten formalen Konsensusverfahren (nominaler Gruppenprozess, Konsensuskonferenz, Delphikonferenz) beraten und als S2-Leitlinie verabschiedet. Die Entwicklungsstufe 3 (S3-Leitlinie) schließlich umfasst alle Elemente einer systematischen Leitlinienentwicklung. Die Definition der AWMF und ÄZQ für eine solche umfassende Leitlinie (s.o.) beinhaltet das Kriterium „systematisch“, das explizit in Abhängigkeit der folgenden fünf, zu erfüllenden Komponenten definiert wird⁸⁸:

1. Logik.
2. Konsens.
3. Evidenzbasierung.
4. EA.
5. Outcomeanalyse.

Nach diesen Empfehlungen sollen mittels einer systematischen EA (z.B. probabilistische Entscheidungsbäume) der erwartete medizinische Nutzen und die Kosteneffektivität medizinischer Vorgehensweisen bestimmt sowie dazu Sensitivitäts- und Schwellenwertanalysen durchgeführt werden⁸⁷.

Ferner wird festgelegt, dass Leitlinien sich heute nicht mehr der Frage verschließen dürfen, wie kosteneffektiv ihre Feststellungen (Statements) sind. Insbesondere Kosteneffektivitätsstudien und Kosten-Nutzwert-Studien (unter Berücksichtigung von QALY) gehören heute zur Evidenzrecherche. Unter anderem sind die Fragen „Ist die Entwicklung und Implementierung einer geplanten Leitlinie effizient?“ und „Können Maßnahmen mit erwiesenem klinischen Nutzen auch effizient erbracht werden, so dass sie für eine Empfehlung in Leitlinien geeignet sind?“ zu beantworten⁸⁶.

Im Rahmen der Outcomeanalyse sollen dabei der durch den Arzt erhobene objektive Gesundheitsstatus, die Selbstbeurteilung der Lebensqualität durch den Patienten und finale Langzeitendpunkte ("true endpoints") berücksichtigt werden⁸⁷.

Damit ist für den deutschen Kontext eine EA eine notwendige Voraussetzung für die Entwicklung einer systematischen Leitlinie. Eine entscheidungsanalytische Modellierung der in der Leitlinie festgelegten medizinischen Vorgehensweisen ist notwendig, um die Optimierung von medizinischem Nutzen und Kosteneffektivität dieser medizinischen Handlungsalgorithmen zu gewährleisten.

3.5.3 Zentrale Versorgungsbereiche und Parameter bei der Evidenzübertragung

Die zweite Forschungsfrage dieses HTA-Berichts betrifft, die Parameter, die in entscheidungsanalytischen Modellierungen zu Übertragung von Evidenz zu berücksichtigen sind. Diese werden im Folgenden systematisch beschrieben.

In vielen Fällen können Ergebnisse einer im Ausland durchgeführten Studie nicht direkt übertragen werden. Um eine Anpassung zu ermöglichen, müssen evtl. aggregierte Parameter zunächst in disaggregierte Parameter zerlegt werden, um dann die für das Zielland erwünschten aggregierten Parameter neu zu berechnen.

Beispielsweise kann bei Screeningprogrammen die Häufigkeit richtig positiver Fälle (aggregierter Parameter) nicht direkt übertragen werden, sondern ist eine Funktion der Größen Prävalenz und Sensitivität (disaggregierte Parameter). Unter bestimmten Umständen kann zwar von einer Übertragbarkeit der Sensitivität ausgegangen werden, aber es muss unter Verwendung der Prävalenz des Ziellands, die gegebenenfalls aus epidemiologischen Studien entnommen werden kann, die für das Zielland valide Häufigkeit richtig positiver Fälle neu berechnet werden. Im Gegensatz dazu spielt die Prävalenz der Erkrankung bei der Beurteilung der Effektivität in therapeutischen Studien meist keine Rolle, da bei Einschluss in die Studienpopulation bereits das Vorliegen der Krankheit vorausgesetzt wird. Ferner können Kosten (aggregierte Parameter) als Produkt verbrauchter Ressourcen (Material, Personal etc.) und Preise ausgedrückt werden. Unterscheiden sich im Zielland nur die Preise, aber nicht der

Ressourcenverbrauch, so können unter Verwendung der Preise des Ziellands die Kostenwerte adaptiert werden.

Diese Tatsache führt zu der Frage, welche Parameter bei der Evidenzübertragung eine Rolle spielen. Anhand einer solchen Liste kann auch detailliert und nachvollziehbar geprüft werden, ob

- sich die Situation im Zielland bezüglich der relevanten Parameter nicht unterscheidet und somit die Studienergebnisse direkt übertragen werden können.
- sich die Situation im Zielland bezüglich relevanter Parameter unterscheidet, die bei einer Übertragung ins Zielland angepasst bzw. substituiert werden müssen.
- die Übertragbarkeit für auch im Zielland nicht verfügbare Parameter in Frage gestellt ist und nur ein neues RCT im Zielland valide Ergebnisse liefern kann.

Ist eine Parameterzerlegung erforderlich, so kann eine Parameterübertragung bzw. -anpassung nur erfolgreich durchgeführt werden, wenn alle für die Übertragung notwendigen Parameter in disaggregierter Form vorliegen. Eine Checkliste für relevante Parameter erfüllt somit auch den Zweck zur Prüfung der Parameter, die aus den Originalpublikationen zu extrahieren bzw. - falls nicht mitgeteilt - von den Autoren der Originalpublikation anzufordern sind.

In Abhängigkeit von Interessenslage und Perspektive der Entscheidungsträger im Gesundheitsbereich ergeben sich unterschiedliche Schwerpunkte von Bereichen, aus denen eine Übertragung von Evidenz wünschenswert wäre. So steht für Ärzte therapeutische Effektivität stärker im Vordergrund als die Kosteneffektivität, die für die Kostenträger jedoch von großer Bedeutung ist.

Im Folgenden werden die Anwendungsbereiche und die einzelnen Parameter(-gruppen) so weit wie möglich generisch zusammengestellt. Diese Listen erheben keinen Anspruch auf Vollständigkeit und sind je nach Fragestellung zu modifizieren.

3.5.3.1 Bereiche in der Versorgungsforschung

Für die Übertragung von Evidenz sind nahezu alle Bereiche im Versorgungswesen von Interesse wie z.B. die Evaluation der Wirksamkeit in der Primärprävention / Reduktion von Risikofaktoren, Früherkennung, Diagnostik, Therapie oder Nachsorge.

Vor dem Hintergrund knapper Ressourcen im Gesundheitsbereich richtet sich das Interesse in den letzten Jahren auch auf gesundheitsökonomische Evaluationen von Medizintechnologien.

In diesem Zusammenhang sind auch die Entwicklung und die Implementierung von Richt- und Leitlinien anzusprechen. Sie existieren in vielen Ländern in zunehmend unüberschaubarer Anzahl und in unterschiedlicher Qualität. Hier könnte eine Übertragbarkeit bzw. Anpassung auf nationale Gegebenheiten zu einer wesentlichen Effizienz- und Qualitätssteigerung führen.

Ver mehrt ins Interesse rücken ferner Aspekte der Versorgungsforschung und der Effizienz von Strukturen des Gesundheitssystems. So werden z.B. kürzlich Unterschiede in der Versorgung von AIDS-Patienten zwischen europäischen Ländern untersucht⁵⁶.

Nicht zuletzt von Interesse sind bestimmte Aspekte der Epidemiologie, insbesondere bei Infektionskrankheiten, die meist große regionale Unterschiede aufweisen⁸⁸.

3.5.3.2 Parameter

Vor der Durchführung der Evidenzübertragung bzw. -anpassung sind die einzelnen Parameter daraufhin zu prüfen, ob sie

- direkt übertragen werden können (dies entspricht formal einer Gleichheitsannahme).
- angepasst bzw. modifiziert werden müssen.
- durch für das Zielland vorliegende Parameter substituiert werden müssen.
- im Zielland nicht vorliegen und dort neu erhoben werden müssen.
- eine Übertragung zu komplex und umfangreich ist und nicht durchgeführt werden kann (in diesem Fall müssten die kompletten Studien im Kontext des Ziellands neu durchgeführt werden).

Die übertragenen, angepassten bzw. substituierten Parameter können schließlich innerhalb eines entscheidungsanalytischen Modells zusammengeführt werden.

Die einzelnen Parameter weisen je nach Kontext und Fragestellung eine unterschiedliche Relevanz auf. So kann beispielsweise die Erkrankungsprävalenz für die Evaluation der Effektivität eines therapeutischen Verfahrens ohne Relevanz sein, jedoch für die anfallenden Gesamtkosten für das Gesundheitssystem einen entscheidenden Faktor darstellen. So wird beispielsweise im Rahmen von entscheidungsanalytischen Prädiktionen darauf hingewiesen, dass ein Massenscreening mit dem PSA-Test für Prostatakarzinom einen bedeutsamen Anteil der Gesamtressourcen des Gesundheitssystems verbrauchen würde³⁹. Im Gegensatz hierzu spielt die Prävalenz im Rahmen der Evaluation diagnostischer oder Screeningverfahren bereits eine entscheidende Rolle bei Bewertung der klinischen Effektivität - unabhängig von ökonomischen Aspekten.

In Tabelle 9 werden die wichtigen Parameter im Rahmen der Evidenzübertragung zusammengestellt. Es handelt sich dabei nicht um eine nur in dieser Form formulierbare und erschöpfende Liste, sondern um eine möglichst umfangreiche Zusammenstellung eines Parametergerüsts, das bei der Evidenzübertragung als unterstützendes Instrument hilfreich sein kann. Die Liste ist bei Bedarf in Abhängigkeit der jeweils spezifischen HTA-Fragestellung zu modifizieren oder zu erweitern.

Die Liste wird unter Berücksichtigung der Arbeiten von Welte et al.⁷⁷ unter Hinzuziehung von Inhalten aus Baltussen et al.² sowie aus Willke et al.⁷⁸ erstellt und lehnt sich teilweise an deren Aufbau an. Ferner werden die Werke von Drummond et al.^{23, 24}, Schöffski et al.⁶¹ und Leidl¹⁹ berücksichtigt. Darüber hinaus werden weitere, unter entscheidungsanalytischen Gesichtspunkten relevante Parameter und Kriterien integriert.

Tabelle 9: Wichtige Parameter bei der Übertragung von Evidenz zwischen verschiedenen Ländern und Versorgungssystemen.

Soziodemographische Parameter

Altersstruktur

Geschlechterzusammensetzung

Ethnische / kulturelle Zusammensetzung, Ausländerquote

Migrationstrends, Reiseaktivität

Gesamt-Hintergrundmortalität / statistische Restlebenserwartung

Geburtenziffern, Alter bei Geburt

Soziale Schichtvariablen (u.a. Verteilung von Bildung, Einkommen)

Fortsetzung Tabelle 9

Epidemiologische Parameter

Prävalenz der Erkrankung(en)

Inzidenz der Erkrankung(en)

Krankheitsspezifische Mortalität (Letalität)

Komorbidität

Verteilung von Risikofaktoren (Genetik, Umweltfaktoren, Ernährungsgewohnheiten, Lebensstile, Gewohnheiten)

Verteilung des Schweregrads

Verteilung von Subtypen / Eigenschaften von Krankheitserregern

Genetische Subformen der Erkrankung

Pharmakogenetische Ausprägungen

Medizinische Outcomeparameter

Präventive / therapeutische / Nachsorgeeffektivität der Technologie (z.B. in Form von Erfolgswahrscheinlichkeiten)

Diagnostische Effektivität der Technologie (z.B. Sensitivität, Spezifität, „Likelihoods“, ROC-Kurven)

Unerwartete Ereignisse der Technologie (Nebenwirkungen, Komplikationen)

Compliance

Patientenpräferenzen / Lebensqualitäts-Outcomeparameter

Lebensqualität / Nutzwerte der verschiedenen relevanten Gesundheitszustände

Kurzzeitveränderungen der Lebensqualität / Nutzwerte während des Einsatzes der Technologie (z.B. Nebenwirkungen)

Ökonomische Parameter

Mengengerüste / Ressourcenverbrauch bei Einsatz der Technologie (einschl. Materialverbrauch, Arzneimittelverbrauch, Personalaufwand etc.)

Marktpreise (u.a. Arzneimittelpreise, Preise für Material und medizinische Geräte etc.)

Gebühren für Versorgungsleistungen (z.B. Fallpauschalen, Abrechnungsgebühren)

Absolute / relative Preise / Gebühren

Indirekte Kosten / Produktivitätsausfallkosten

Externe Effekte

Skaleneffekte, Kapazitätsausnutzung

Währungskonvertierungen

Referenzjahre, Indexierung, Inflation

Diskontrate

Zuzahlungen

Positiv- / Negativlisten

Marktanteile von Generika

Fortsetzung Tabelle 9

Versorgungs- und Technologiestandards

Nationale / regionale Leitlinien und Versorgungsstandards, ärztliche Präferenzen und Gewohnheiten

Verfügbarkeit von Technologien, unterschiedliche Einführungsphasen, unterschiedliches Angebot bzw. fehlende Technologien / Produkte

Unterschiedliche Varianten der Technologien

Schwierigkeit des Verfahrens, Lernkurven bei neuen Technologien

Vergleichstechnologie (derzeitiger Routinestandard)

Art und Anzahl von Alternativtechnologien

Strukturelle Parameter des Gesundheitssystems

Ressourcenverteilung im Gesundheitssystem

Schnittstellen der Versorgungssektoren

Rahmenbedingungen für Technologien

Aufgabenverteilung zwischen ärztlichem und Pflege- oder medizinischem Hilfspersonal

Zugänglichkeit (Gatekeeper-Prinzip)

Wissensstand und Qualifikation, Produktivität und Gehaltsstruktur des Personals

Anreizstrukturen (stationäres Honorierungsmodell, angebotsinduzierte Nachfrage, Moral Hazard)

Weitere Determinanten

Rechtliche Rahmenbedingungen

Ethische und kulturelle Prämissen

Zu ergänzenden Ausführungen bezüglich der Rolle einzelner Parameter sei auf die Literatur verwiesen^{3, 24, 32, 34, 61, 77}.

3.5.3.2.1 Soziodemographische Parameter

Von grundlegender Bedeutung sind die Auswirkungen soziodemographischer Parameter. Die Altersstruktur der Studienpopulation deren Daten übertragen werden sollen, kann weitere wesentliche Größen wie Erkrankungsrisiko, Heilungschancen, Häufigkeit von Komplikation, die Restlebenserwartung, Kostenstrukturen oder auch die Produktivität stark beeinflussen. Effekt- und Kostenseite können durch eine unterschiedliche Geschlechterzusammensetzung verzerrt werden, da das Geschlecht Risikofaktor für eine Reihe von Krankheiten ist, so z.B. die Magersucht oder die Multiple Sklerose⁵⁰. Die Gesamtmortalität (bzw. statistische Restlebenserwartung) oder die krankheitsspezifische Mortalität (Letalität) können besonders bei Aussagen über lebensverlängernde Therapien wesentliche Verzerrungsgrößen darstellen. Demographische Kenngrößen wie die Geburtenziffer oder das durchschnittliche Alter bei der Geburt eines Kindes können gelegentlich von Bedeutung sein. So geht ein höheres Durchschnittsalter bei der Geburt mit einer höheren Rate an Komplikationen in der Schwangerschaft oder genetischen Defekten des Kindes einher¹⁴. Migrationstrends, Ausländerquote oder auch Reiseaktivität können sich in Veränderungen der Krankheitslast oder des Krankheitsspektrums widerspiegeln. So werden z.B. Infektionen mit Hepatitis A in der deutschen Bevölkerung im Wesentlichen bei Reisen erworben⁸⁹. Variablen der sozialen Schicht wie der Verteilung von

Bildungsgraden oder Einkommen sind häufig mit der Verteilung von Risikofaktoren bestimmter Erkrankungen assoziiert (z.B. Rauchen und koronare Herzerkrankung).

3.5.3.2.2 Epidemiologische Parameter

Zentrale epidemiologische Kenndaten determinieren fast immer die Übertragbarkeit von Ergebnissen. Ohne Anpassungen durch die entsprechenden Werte im Zielkontext ist üblicherweise mit großen Verzerrungen zu rechnen. Zu den wichtigsten epidemiologischen Parametern gehören neben der Prävalenz und Inzidenz der interessierenden Erkrankung die krankheitsspezifische Mortalität, die Verteilung der Komorbidität (z.B. Diabetes bei koronarer Herzkrankheit), die dazugehörige Verteilung von Risikofaktoren und die Verteilung des Krankheitschweregrads. Mitunter spielen genetische Faktoren eine Rolle für die Progression der Erkrankung oder die Ansprechrate auf Arzneimittel.

Risikofaktoren können Umwelteinflüsse, Lebensstile, Ernährungsgewohnheiten, Einstellungen und genetische Dispositionen (z.B. Acetyliererstatus bei kolorektalen Karzinomen) umfassen. So ist offensichtlich, dass bei der Übertragung von Evidenz zur koronaren Herzkrankheit zwischen verschiedenen Ländern deren Prävalenz durch den Anteil der Raucher mit beeinflusst wird. Unterschiedliche Ernährungsgewohnheiten können über verschiedenes Präventionspotential ebenfalls zu diversen Krankheitsprävalenzen führen. Speziell bei Infektionskrankheiten kann die Verteilung von Eigenschaften und Subtypen der Erreger eine große Rolle spielen. Beispiele hierfür sind die geographisch verschieden verteilten Genotypen des Hepatitis C Virus³⁷, die verschiedenen Formen der Malaria und das hämolytisch-urämische Syndrom, das nur durch bestimmte Stämme von *E. coli* ausgelöst wird⁵⁰. Die jüngsten Entwicklungen im Bereich der Pharmakogenomik bieten das Potential zur Berücksichtigung von genetisch determinierten Arzneimittelansprechraten und damit einer Reduzierung des so genannten „Pharmacogenomics-Bias“⁶⁶.

3.5.3.2.3 Medizinische Outcomeparameter

Medizinische Outcomeparameter werden häufig ohne Adjustierungen zwischen Ländern übertragen und damit implizit gleichgesetzt. Dies ist für viele, nicht genetisch oder extern bedingten Ansprechraten von Technologien durchaus sinnvoll. Unterscheiden sich aber Compliance oder Pharmakogenomik zwischen den Ländern, ist eine direkte Übertragung nicht möglich, sondern es muss eine Adjustierung mit den disaggregierten Daten (z.B. Erfolgsraten für compliante / nicht compliante Patienten oder die verschiedenen genetischen Ausprägungen und deren Häufigkeiten im Zielland) vorgenommen werden. Sind diese für das Zielland nicht bekannt, muss die Unsicherheit in Sensitivitätsanalysen berücksichtigt werden.

Insbesondere bei der Übertragung der aus RCT ermittelten Wirksamkeit (Efficacy) einer Technologie auf Verhältnisse unter Routinebedingungen (Effectiveness) sind unterschiedliche Raten unerwarteter Ereignisse (wie Nebenwirkungen oder Komplikationen) und Compliance sowohl bei klinischen als auch bei gesundheitsökonomischen Evaluationen zu berücksichtigen.

Bei der Übertragung von Evidenz zu diagnostischen Verfahren ist zu prüfen, ob die Testcharakteristika wie Sensitivität und Spezifität ausreichend robust sind. Prädiktive Werte sind aufgrund der Abhängigkeit von der Erkrankungsprävalenz neu zu bestimmen. Insbesondere bei Screeningverfahren kommt es häufig vor, dass die Testgüteparameter Sensitivität und Spezifität verschiedener Testverfahren aus Gründen der statistischen Effizienz an vorselektierten Populationen mit erhöhtem Krankheitsrisiko ermittelt werden. Für den Fall, dass eine erhöhte Krankheitsprävalenz auch mit fortgeschrittenen Krankheitsstadien einhergeht (z.B. Krebserkrankungen), stellen die anhand von Risikopopulationen ermittelten Werte für Sensitivität Überschätzungen dar (Casemix-Phänomen). Werden diese Werte für die Evaluation von bevöl-

kerungsweiten Screeningprogrammen verwendet, so kann es zu einer Überschätzung des klinischen Nutzens des Screeningprogramms kommen. Ein Beispiel hierfür ist das Screening für Zervixkarzinom mit dem Papanicolaou-Test. Der Vergleich der Ergebnisse zu Sensitivität und Spezifität des Papanicolaou-Tests mit und ohne Ausschluss von Studien an Risikopopulationen zeigt eine deutliche Reduzierung der Sensitivität von 69 auf 47 % nach Ausschluss von Studien an Risikopopulationen⁵².

3.5.3.2.4 Patientenpräferenzen / Lebensqualitätsoutcomeparameter

Die Prüfung der Übertragbarkeit von Patientenpräferenzen spielt insbesondere bei Kosteneffektivitätsstudien von Technologien eine Rolle, deren klinischer Nutzen weniger auf einem lebensverlängernden als vielmehr auf einem lebensqualitätserhöhenden Effekt beruht (z.B. Dorsopathie-Präventionsprogramme). In solchen Analysen besteht der Nenner des Kosteneffektivitätsverhältnisses aus der inkrementellen qualitätskorrigierten Lebenserwartung (QALY), die wiederum direkt von den Nutzwerten (Utilities) der einzelnen Gesundheitszustände abhängt²⁸. Da die Arbeitsfähigkeit meist als Komponente in die Erfassung der gesundheitsbezogenen Lebensqualität eingeht, kann davon ausgegangen werden, dass Gesundheitszustände mit eingeschränkter Arbeitsfähigkeit in Ländern mit einem höheren Sozialversicherungsanteil mit einem geringeren Lebensqualitätsverlust bewertet werden, als in Ländern mit einem höheren Anteil an schlechter oder überhaupt nicht gegen Arbeits- / Berufsunfähigkeit versicherten Patienten. Ferner können Patientenpräferenzen aufgrund unterschiedlicher kultureller oder ethischer Hintergründe abweichen. Wenn im Zielland keine Nutzwerte vorliegen, müssen diese entweder neu erhoben oder aber die systematische Unter- oder Überschätzung in Sensitivitätsanalysen berücksichtigt werden.

3.5.3.2.5 Ökonomische Parameter

Gesundheitsökonomische Studienergebnisse z.B. aus Kosteneffektivitätsanalysen sind selten direkt zu übertragen. Selbst wenn von einer Gleichheit der medizinischen Effektivität im Zielland ausgegangen werden kann, kann dies in den seltensten Fällen für die Kosten angenommen werden. Kosten sind das Produkt aus Mengen und Preisen, deshalb sind bei der Übertragung idealerweise die Mengen- und Preisgerüste zu disaggregieren. Preisgerüste unterscheiden sich nahezu regelmäßig zwischen verschiedenen Ländern, Mengengerüste in den meisten Fällen. Diese Unterschiede sind für einen hohen Anteil der Variabilität von Kosten-Nutzen-Verhältnissen verantwortlich⁷⁸. Bei den Preisen ist auf die Mechanismen der Preisbildung (echte Marktpreise, Schattenpreise oder festgelegte Gebühren), auf die Berechnung als absolute oder relative Preise sowie gegebenenfalls auf Währungskonvertierungen und Referenzjahre (gegebenenfalls Indexierung zur Berücksichtigung von Inflation) zu achten, da Änderungen der Kosten über die Zeit möglich sind. Gegebenenfalls ist die Diskontierung an die länderspezifischen Empfehlungen anzupassen. Es besteht gemeinhin Konsens darüber, dass Preise nicht übertragen, sondern jeweils spezifisch für das Zielland ermittelt werden sollten. Die einfachste Variante der Übertragung von Kosteneffektivitätsverhältnissen bei im Ausland durchgeführten gesundheitsökonomischen Evaluationen besteht also in der Übernahme des Mengengerüsts und der Substitution des ziellandspezifischen Preisgerüsts. Für die Übertragbarkeit der Mengengerüste, deren Ermittlung sich oft sehr aufwendig gestaltet, bestehen unterschiedliche Ansichten, wenngleich die spezifische Erhebung als valider anzusehen ist⁶¹. Allerdings können Mengengerüsten oft strukturbedingt nicht direkt übertragen werden. So werden beispielsweise in den USA häufig die Anzahl der Arztkontakte erfasst, die in Deutschland aufgrund hiesiger Abrechnungsgegebenheiten nur eine untergeordnete Rolle spielen.

Werden indirekte Kosten in die Analyse eingeschlossen, so sind diese grundsätzlich spezifisch für das Zielland festzulegen. Für Deutschland wird beispielsweise die Verwendung eines

durchschnittlichen Einkommens für die Berechnung von Produktivitätsausfallskosten vorgeschlagen¹.

Als weiteres Kriterium, das allerdings nicht für eine Übertragung selbst in Frage kommt, sondern bei Diskordanz eher die Übertragbarkeit als solche in Frage stellt, ist die Wahl der Perspektive der gesundheitsökonomischen Evaluation.

Die Preisbildung im Arzneimittelbereich wird aufgrund ihrer herausragenden Bedeutung für die Kosten eines Gesundheitssystems von Regierungen meist besonders reguliert. So unterscheiden sich beispielsweise Zuzahlungsregelungen durch Patienten, Kostenerstattungen und Verordnungsgewohnheiten z.B. aufgrund von Negativisten oder des Marktanteils günstigerer Generikaprodukte (z.B. aufgrund politischer Erwünschtheit)³⁴. Die Relevanz der zuletzt genannten Parameter ist in besonderer Weise von der Wahl der Perspektive gesundheitsökonomischer Studien abhängig. So haben unterschiedliche Zuzahlungsregelungen für eine Beurteilung aus der Sicht von Krankenkassen eine andere Bedeutung als bei einer gesamtgesellschaftlichen Perspektive. Neben dem jeweiligen gesundheitspolitischen Stellenwert einer Kostendämpfungspolitik können bei entsprechender nationaler Bedeutung der pharmazeutischen Industrie im Rahmen arbeitsmarktpolitischer Erwägungen protektionistische Tendenzen wohl erwartet werden³⁴.

Die quantitative Nachfrage nach Ressourcen kann über Skaleneffekte und Aspekte der Kapazitätsausnutzung zu einem anderen Preisgefüge führen. Skaleneffekte bezeichnen das Verhältnis von Output- zu Inputsteigerungen. Der Grad der Kapazitätsausnutzung wirkt sich über die Effizienz der Produktion auf Kosten und Preise aus, kann aber im Rahmen einer angebotsinduzierten Nachfrage auch das Mengengefüge beeinflussen. So steht z.B. zu erwarten, dass bei zwischen zwei Ländern deutlich unterschiedlicher Versorgungsquote von Kernspintomographen sich auch deren Inanspruchnahme unterscheiden wird.

3.5.3.2.6 Versorgungs- und Technologiestandards

Unterschiedliche Leitlinien und Praxisstandards können zu verschiedenen Ergebnissen sowohl auf der klinischen als auch auf der ökonomischen Seite führen. Dies betrifft nicht nur die Standards zum Einsatz und der Durchführung der untersuchten Technologie, sondern auch die Standards zur Versorgung von Folgeerkrankungen. Beispielsweise kommt der frühzeitigen Therapie von Organerkrankungen ein größerer inkrementeller Nutzen zu, wenn Organtransplantate nicht oder nicht in ausreichender Anzahl zur Verfügung stehen.

Bei der Beurteilung der Wirksamkeit von Verfahren unter Routinebedingungen im Zielland sind Unterschiede sowohl bei der Übertragung RCT-Routineverhältnisse als auch bei der Übertragung Ursprungsland-Zielland zu berücksichtigen. Hierbei spielen ärztliche Präferenzen sowie regionale Therapiestandards und -gewohnheiten eine Rolle. So kann sich beispielsweise eine unterschiedliche Hygienepaxis auf die Komplikationsrate und damit auf die Kosten und das medizinische Ergebnis auswirken⁷⁷.

Die Verfügbarkeit von Technologien zur Behandlung der interessierenden Erkrankung oder ihrer Folgeerkrankungen kann das gesamte Ergebnis eines HTA-Berichts verändern. Da diese Gegebenheiten nicht qualitative, sondern strukturelle Determinanten darstellen, sind sie für eine valide Evaluation in der Struktur des entscheidungsanalytischen Modells zu berücksichtigen. Beim bereits vorliegenden entscheidungsanalytischen Modell ist dies entsprechend strukturell zu adaptieren.

Es ist zu prüfen, ob die Technologie im Zielland ebenso durchgeführt wird wie in der Studie aus der die zu übertragenden Daten stammen. Eine andere Variante der Technologie ist im

Rahmen von HTA streng genommen als andere Technologie zu bezeichnen und zu bewerten.

Bei neu eingeführten Technologien sollte auf möglicherweise unterschiedliche Lernkurven in der Anfangsphase geachtet werden; die Lernkurven selbst sind wiederum abhängig von der Schwierigkeit des neuen Verfahrens.

Die Vergleichstechnologie sowie die Art und Anzahl der zur Verfügung stehenden Alternativtechnologien bestimmen sowohl den inkrementellen medizinischen Nutzen als auch das inkrementelle Kosten-Nutzwert-Verhältnis der untersuchten Technologie. Aus diesem Grund sind die für das Zielland spezifische Vergleichs- und Alternativtechnologien ins entscheidungsanalytische Modell zu integrieren.

3.5.3.2.7 Strukturelle Parameter des Gesundheitssystems

Es kann davon ausgegangen werden, dass Anpassungen bei der Übertragung in unterschiedlich strukturierte Gesundheitssysteme mit besonders komplexen Modellierungen einher zu gehen haben. So ist auf möglicherweise anders verteilte Ressourcen im Gesundheitssystem abzustellen und die sich gegebenenfalls hieraus entwickelnde Schnittstellenproblematik zu identifizieren (z.B. unterschiedliche Zuständigkeitsverteilungen für ambulanten und stationären Sektor, abweichende Krankenhausverweildauern⁵⁶).

Zu den strukturellen Eigenheiten ist auch eine möglicherweise unterschiedliche Aufgabenverteilung zwischen ärztlichem und Pflege- oder medizinischem Hilfspersonal zu zählen (z.B. Blutentnahme), Verteilung der Versorgungsstufe medizinischer Infrastruktur (Kreiskrankenhäuser oder Universitätskliniken) sowie deren Zugänglichkeit (z.B. Gatekeeper-Prinzip der Hausärzte in Holland oder Großbritannien). Schließlich können die Rahmenbedingungen für die Durchführbarkeit neuer Technologien so differieren, dass sich eine Übertragung verbietet. So sind Überlegungen zur Einführung eines Screenings aller neugeborenen Kinder auf Schwerhörigkeit so lange obsolet, wie keine entsprechenden Kapazitäten zur Behandlung vorhanden sind. Ebenso ist die Übertragung einer neuen molekularbiologischen Nachweisteknik ohne entsprechend qualifiziertes bzw. geschultes Personal wenig sinnvoll.

Zu den relevanten lokalen Gegebenheiten zählen Wissensstand und Qualifikation des mit einem Verfahren befassten Personals, seine Produktivität und seine Gehaltsstruktur. Auswirkungen können dabei sowohl auf die Kosten, als auch auf die medizinische Effizienz möglich sein.

Durch unterschiedliche Strukturen in den nationalen Gesundheitssystemen können sich verschiedene Anreizstrukturen für die Leistungserbringer ergeben und so diskordante Effekte zwischen Ausgangs- und Zielland hervorrufen. Neben einem unterschiedlichen Ausprägungsgrad angebotsinduzierter Nachfrage, sind verschiedene Praktiken der Kostenerstattung und Ausformungen des Versicherungsschutzes (Moral Hazard) von wesentlicher Bedeutung. So besteht beispielsweise bei pauschalierten Versicherungspaketen vermehrt die Tendenz zu gehäuften Arztkontakten oder übertriebener Arzneimittelversorgung (Medikamente auf Vorrat). Für den Krankenhaussektor zeigen unterschiedliche Vergütungsformen nachhaltige Effekte. So sind die Krankenhausliegezeiten bei Honorierung über Fallpauschalen, wie sie auch in Deutschland eingeführt werden sollen (DRG) deutlich niedriger, als bei tagesgenauer Abrechnung über Pflegesätze⁶¹.

3.5.3.2.8 Weitere Determinanten

Bei der Prüfung zur Übertragbarkeit von Technologien und Evidenz sind selbstverständlich auch rechtliche Rahmenbedingungen in Betracht zu ziehen. So unterscheiden sich die Regelungen zum Datenschutz in den USA und in Deutschland beträchtlich voneinander. Dies

kann Auswirkungen auf die Kostenseite haben.

Häufig werden auch unterschiedliche ethische Prämissen zu berücksichtigen sein. Ein aktuelles Beispiel hierfür ist die derzeitige Diskussion um Forschung an embryonalen Stammzellen, oder auch die Diskussion um die ärztliche Sterbehilfe in Holland.

Ferner ist fraglich, ob die in einem Land akzeptierten Schwellenwerte der gesellschaftlichen Zahlungsbereitschaft (z.B. Schwellenwert für Kosteneffektivitätsverhältnisse) auch in anderen Ländern Akzeptanz finden. Das Verhältnis zwischen Nutzenmaximierung (utilitaristisches Prinzip) und der Verteilungsgerechtigkeit (Gleichheitsprinzip) wird in verschiedenen Staaten oder Kulturkreisen unterschiedlich ausgerichtet sein. Ebenso sind Fragen der Selbstbestimmungsrechte (Patientenrechte) und Eigenverantwortlichkeiten von Patienten kulturspezifisch. Die Auswirkung derartiger Unterschiede auf der Kosten- oder Effektseite sind allerdings kaum quantitativ zu erfassen und müssen deshalb beim Übertragbarkeitsprozess qualitativ diskutiert und beurteilt werden.

Zu ergänzenden Ausführungen bezüglich der Rolle einzelner Parameter sei auf die Literatur verwiesen^{3,24, 34, 61, 77}.

3.5.4 Entscheidungsanalyse als Instrument zur Evidenzübertragung

Im Folgenden wird Bezug auf die dritte und vierte Forschungsfrage genommen. Es wird dargelegt, welche strategischen Vorgehensweisen zur Evidenzsynthese und -übertragung im Rahmen deutscher HTA-Projekte unter gegebenen Rahmenbedingungen sinnvoll sind und welche Möglichkeiten und Limitationen die EA dabei besitzt.

Die Prüfung und Durchführung der Übertragung von internationaler Evidenz auf den Kontext eines bestimmten Lands kann anhand eines vereinfachten Schemas erläutert werden, das auf den Komponenten Entscheidung, Evidenz und Modell basiert. Das Modell besteht dabei aus einer strukturellen (Modellstruktur) und einer numerischen Komponente (Modellparameter). Das Schema findet sich in Abbildung 19.

Wie bereits weiter oben dargelegt wird, kann selten ohne weitere Prüfung davon ausgegangen werden, dass evidenzbasierte Entscheidungen, die in einem bestimmten Ursprungsland (Land A) getroffen werden, direkt auf das Zielland (Land B) übertragbar sind.

Ist die Entscheidung in Land A evidenzbasiert, so kann gegebenenfalls die Evidenzgrundlage (Evidenz A) auf das Zielland übertragen werden. Dies bedeutet nicht immer automatisch, dass auch die Entscheidung im Zielland (Entscheidung B) dieselbe sein muss wie im Ursprungsland (Land A), da unterschiedliche Rahmenbedingungen in der Zielpopulation und im Gesundheitssystem existieren können (z.B. die Erkrankungsprävalenz, die Verfügbarkeit weiterer in die Analyse einzubeziehende Alternativtherapien oder eine unterschiedliche gesellschaftliche Zahlungsbereitschaft). Kann allerdings das gesamte Ergebnis der ausländischen Studie übertragen werden, kann in vielen Fällen eine qualitative Beurteilung ausreichen, um von dieser Evidenz ausgehend zu einer Entscheidung im Zielland zu kommen. In manchen Fällen können dazu einfache quantitative Rechenschritte notwendig sein, die als einfachste Form der formalen EA betrachtet werden können.

In den meisten Fällen wird man jedoch anhand der oben dargestellten Parameterliste (Tabelle 9) feststellen, dass sich manche numerischen Parameter oder strukturellen Rahmenkriterien beim Vergleich der Länder unterscheiden. Dies bedeutet formal, dass für beide Länder verschiedene Modelle gelten. In diesem Fall ist eine Zerlegung der aggregierten numerischen Ergebnisse der Studie im Ursprungsland (Land A) in die disaggregierten numerischen Parameter erforderlich. Es ist dann zu bestimmen, welche der numerischen Parameter komplett

substituiert werden müssen, welche durch mathematische Standardisierung oder Adjustierung an das Zielland angepasst werden können sowie bei welchen davon ausgegangen werden kann, dass sie sich nicht zwischen den Ländern unterscheiden und somit direkt übertragen werden können. Sind Werte komplett zu substituieren, so muss entweder auf vorliegende Daten aus ziellandspezifischen Studien oder Datenbanken zurückgegriffen, oder – falls solche fehlen – müssen neue Primärerhebungen für diese Parameter durchgeführt werden. Ist letzteres nicht möglich, scheitern der formale sowie der transparente quantitative Prozess und es muss auf qualitative Verfahren wie Expertenmeinungen zurückgegriffen werden.

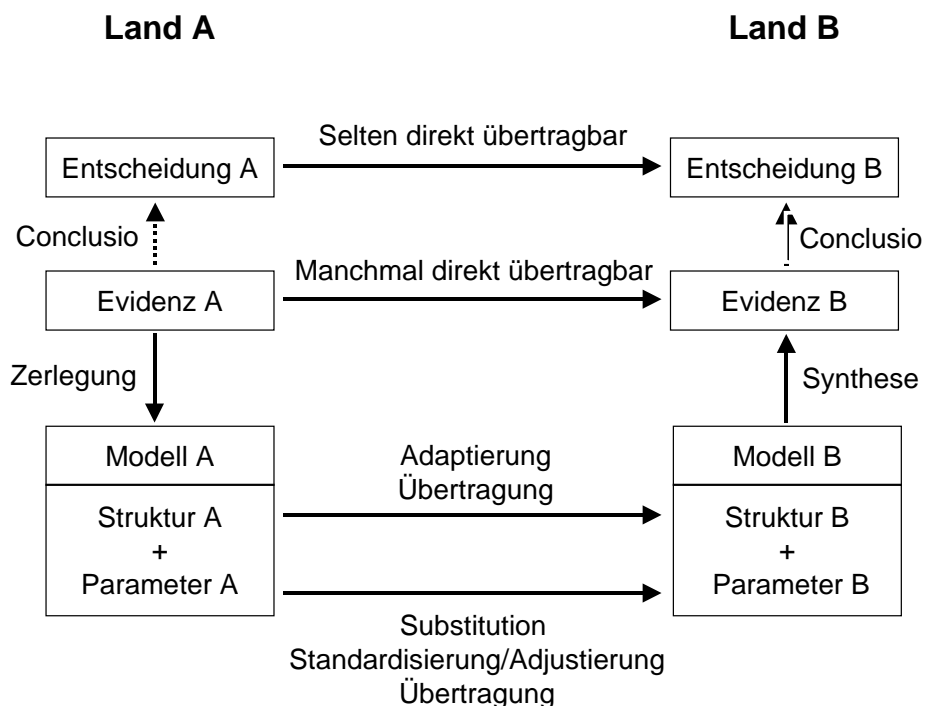
Gibt es strukturelle Unterschiede in den Gesundheitssystemen der Länder, sind diese Unterschiede bei der Entwicklung des Modells im Zielland zu berücksichtigen und in Form von strukturellen Adaptierungen zu implementieren.

Liegt ein explizites entscheidungsanalytisches Modell für das Ursprungsland (Modell A) vor, kann dieses nach entsprechender Modifikation für das Zielland (Modell B) verwendet werden. Ist ein solches Modell noch nicht veröffentlicht oder existieren die Modelle der Ursprungsländer nur implizit, muss für den Kontext des Ziellands ein entscheidungsanalytisches Modell neu entwickelt, formuliert und programmiert werden.

Die Einbettung der für das Zielland gültigen Parameterschar (Parameter B) in die Modellstruktur des Ziellands (Struktur B) macht das entscheidungsanalytische Modell des Ziellands aus (Modell B). Dieses Modell kann nun verwendet werden, um die einzelnen disaggregierten Parameter zu synthetisieren und die interessierende Evidenz in Form aggregierter Zielgrößen zu ermitteln (Evidenz B).

Die ziellandspezifische Evidenz (Evidenz B) ist Grundlage für die Schlussfolgerungen und Empfehlungen an die Entscheidungsträger im Zielland (Entscheidung B).

Abbildung 19: Schematische Darstellung des Vorgehens bei der Evidenzübertragung unter Verwendung der Komponenten Entscheidung, Evidenz und Modell.



Basierend auf diesem schematischen Gerüst werden im nächsten Abschnitt die konkreten und praktischen Schritte in Form eines Stufenkonzepts zur Evidenzübertragung beschrieben.

3.5.5 Stufenkonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext

Tabelle 10 gibt eine Übersicht über die notwendigen Schritte zur Übertragung von internationaler Evidenz auf den nationalen Kontext. Anschließend werden die einzelnen Schritte des Stufenkonzepts ausführlich erläutert.

Tabelle 10: Stufenkonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext.

1. Identifikation internationaler Evidenz.
 2. Identifikation der am Entscheidungsproblem beteiligten und relevanten Parameter.
 3. Bewertung der Übertragbarkeit der internationalen Ergebnisse anhand einer Prüfliste⁷⁷.
 4. Entscheidung, ob eine Reanalyse für den deutschen Kontext notwendig ist.
 5. Entscheidung, ob eine Entscheidungsanalyse sinnvoll / notwendig ist.
 6. Aufbau eines interdisziplinären Expertennetzwerks bestehend aus Experten der Fachgebiete Medizin, Epidemiologie, Gesundheitsökonomie, Lebensqualitätsforschung, Biometrie, Medizinische Informatik, Gesundheitspolitik, Ethik und Entscheidungswissenschaften (Decision Sciences).
 7. Prüfen der Datenverfügbarkeit für Deutschland.
 8. Erhebung von validen landesspezifischen Daten bezüglich Epidemiologie, Krankheitsverlauf, Wirksamkeit, Lebensqualität, Standard klinischer Versorgungspraxis, Einsatz der untersuchten Technologie / Vergleichstechnologie / Alternativtechnologie, Ressourcenverbräuche und Preise.
 9. Ergänzen der Daten durch Expertenschätzungen.
 10. Modellentwicklung oder strukturelle Modelladaptation.
 - Identifikation von publizierten internationalen medizinischen oder gesundheitsökonomischen Entscheidungsanalysemodellen.
 - Systematische Bewertung der Modelle und Selektion von einem oder mehreren qualitativen hochwertigen entscheidungsanalytischen Modellen
 - Kommunikation mit den Modellautor und Aufbau einer Kollaboration.
 - Evaluation der Transparenz, Zugänglichkeit, Flexibilität und Adaptationsmöglichkeit der selektierten Modelle.
 - Falls keine internationalen entscheidungsanalytischen Modelle publiziert sind, prüfen der Entwicklung und Durchführbarkeit einer eigenen entscheidungsanalytischen Modellierung.
 - Übertragung und Programmierung der internationalen Modellstruktur in anwenderfreundliche Software.
 - Prüfen der Übertragbarkeit der internationalen Modellstruktur und Adaptation der Modellstruktur an die landesspezifische Situation.
 11. Numerische Implementation oder Adaptation der Modellparameter.
 - Prüfen der Übertragbarkeit der internationalen Modellparameter und Adaptation von Modellparametern an die jeweiligen landesspezifischen Situationen: Übergangswahrscheinlichkeiten im Krankheitsverlauf, medizinische Effektparameter, Kostenparameter, Lebensqualitätsparameter.
 12. Klinische und / oder Kosteneffektivitätsanalyse.
 13. Sensitivitätsanalysen für unsichere Parameter oder Parameter aus Expertenschätzungen.
 14. Ergebnisinterpretation und Schlussfolgerungen / Empfehlungen im Gesamtkontext.
 15. Festlegung des weiteren Forschungsbedarfs anhand der entscheidungsrelevanten Parameter, die in Sensitivitätsanalysen oder „Value-of-Information-Analysen“ identifiziert werden.
-

Im Folgenden sollen die Einzelschritte der Übersicht detaillierter beschrieben werden.

3.5.4.1 Identifikation internationaler Evidenz

Eine Identifikation der aktuell publizierten wissenschaftlichen Evidenz zur definierten Fragestellung einschließlich dazu vorliegender entscheidungsanalytischer Modelle sollte durch eine systematische Literaturrecherche in relevanten Datenbanken durch Verwendung gebräuchlicher Schlüsselworte (Keywords) sowie durch Freihandsuche unter Verwendung einer Suchmaschine im Internet stattfinden.

3.5.4.1.1 Identifikation der am Entscheidungsproblem beteiligten und relevanten Parameter

Eine Identifikation der am Entscheidungsproblem beteiligten und relevanten Parameter sollte anhand der Fragestellung und der in der Literatur berichteten Studien zu ähnlichen Problemstellungen sowie unter Verwendung von Expertenangaben zum Krankheitsverlauf, Epidemiologie und medizinischer Versorgungspraxis erfolgen.

3.5.4.1.2 Bewerten der Übertragbarkeit der internationalen Ergebnisse durch Prüflisten

Die Übertragung der internationalen Evidenz sollte systematisch auf Grundlage von Prüflisten erfolgen, um eine objektive Einschätzung zu gewährleisten. Hierzu hat die German Scientific Working Group Technology Assessment for Health Care eine Reihe von Prüflisten zur Evaluation der Übertragbarkeit und der Bewertung der Studienqualität für verschiedene Studienarten im Rahmen der Technologieevaluation in HTA entwickelt⁴³.

3.5.4.1.3 Entscheidung, ob eine Reanalyse für den deutschen Kontext notwendig ist

Unter Verwendung systematischer Prüflisten zur Übertragbarkeit internationaler Evidenz auf den nationalen Kontext sollten eine Synthese der Evaluation und eine Entscheidung darüber fallen, ob die gefundene internationale Evidenz auf den nationalen Kontext übertragbar ist. Ist dies nicht in zufrieden stellender Weise der Fall, sollte über die Notwendigkeit einer Reanalyse für den nationalen Kontext mit entsprechender Erhebung evtl. relevanter Daten für den deutschen Versorgungskontext entschieden werden.

3.5.4.1.4 Entscheidung, ob eine Entscheidungsanalyse sinnvoll / notwendig ist

Die Durchführung einer EA kann dann sinnvoll sein, wenn simultan mehrere relevante Parameter der Problemstellung von der internationalen Evidenz abweichen und gleichzeitig das zu evaluierende Endergebnis in verschiedene Richtungen beeinflussen, so dass eine Abschätzung der Richtung der Abweichung des Ergebnisses unter nationalen Bedingungen schwer durchführbar ist.

3.5.4.1.5 Aufbau eines interdisziplinären Expertennetzwerks

In EA werden Informationen aus verschiedenen Disziplinen zusammengeführt; dabei ist eine detaillierte Kenntnis der fachspezifischen Konzepte und Methoden erforderlich. Beteiligte Disziplinen umfassen Medizin, Epidemiologie, Gesundheitsökonomie, Lebensqualitätsforschung, Biometrie, Medizinische Informatik, Gesundheitspolitik, Ethik und Entscheidungswissenschaften (Decision Sciences).

Da die Public Health-Studiengänge in Deutschland zu den wenigen Ausbildungsgängen gehören, in welchen ein Großteil dieser Gebiete zumindest einführend gelehrt wird, ist es bei einem entscheidungsanalytischen Projekt von großem Vorteil, wenn ein Public Health-Experte oder ein Experte mit ähnlicher Ausbildung und Erfahrung als Mitwirkender oder Berater im HTA-Projekt eingesetzt werden kann. Da kein einzelner Forscher die erforderliche Expertise in allen

beteiligten Disziplinen mitbringen kann, steht und fällt die Durchführbarkeit und Qualität der entscheidungsanalytischen Evidenzübertragung mit dem Einrichten eines interdisziplinären Netzwerks.

3.5.4.1.6 Prüfen der Datenverfügbarkeit für Deutschland

Für die erhobenen relevanten Parameter der Problemstellung sollte die Datenverfügbarkeit für den nationalen Kontext geprüft werden. Hierzu sollte eine Erhebung von validen landesspezifischen Daten bezüglich Epidemiologie, Krankheitsverlauf, Wirksamkeit, Lebensqualität, Standard klinischer Versorgungspraxis, Verbreitung und Einsatz der untersuchten Technologie bzw. Vergleichs- oder Alternativtechnologien, Ressourcenverbräuchen und Preisen erfolgen. Diese können sowohl in Primäruntersuchungen (d.h. eigener Datenerhebung) als auch in Sekundäranalysen (d.h. aus der Literatur) gewonnen werden^{44, 45}. Liegen mehrere Ergebnisse aus der wissenschaftlichen Literatur oder aus medizinischen bzw. gesundheitsökonomischen Datenbanken vor, können Metaanalysen eingesetzt werden, um Maßzahlen mit hohem Evidenzgrad zu erhalten⁷⁵. Nicht immer sind alle erforderlichen epidemiologischen, klinischen und ökonomischen Daten für alle im Modell als relevant aufgenommenen Gesundheitszustände oder Übergangswege präsent. In diesen Fällen müssen in interdisziplinärer Zusammenarbeit der verschiedenen Experten entweder zusätzliche numerische Annahmen bezüglich der Modellparameter gemacht werden oder die Modellstruktur muss vereinfacht werden. Letzteres impliziert ebenfalls bestimmte Annahmen, die explizit zu formulieren, von den Experten auf Plausibilität zu prüfen und von den Analysten in Sensitivitätsanalysen zu verändern sind. Eine systematische Zusammenstellung der erhobenen Daten sollte mittels Erstellung von Datenbanken für die medizinische Evidenz, Ressourcenverbräuche und Kosten, sowie für Lebensqualitätsdaten durchgeführt werden.

3.5.4.1.7 Modellentwicklung oder strukturelle Modelladaptation

Ist die Entscheidung für eine Durchführung einer EA getroffen worden, kann dieses entweder durch die eigenständige Entwicklung eines Modells oder durch die Anpassung und Weiterentwicklung bereits publizierter internationaler Modelle geschehen. In beiden Fällen sollte als erster Schritt eine Identifikation von publizierten internationalen medizinischen oder gesundheitsökonomischen EA-Modellen durch eine entsprechend der Problemstellung ausgerichtete Literaturrecherche stehen. Eine systematische Bewertung der identifizierten Modelle anhand von systematischen Prüflisten eignet sich zur objektiven Evaluation und Selektion qualitativ hochwertiger entscheidungsanalytischer Modelle.

Sind zu der betreffenden Problemstellung bisher keine entscheidungsanalytischen Modelle publiziert, empfiehlt es sich nach entscheidungsanalytischen Modellen zu suchen, die zumindest die Zielerkrankung betreffen. Es ist dann abzuwägen, ob eine eigenständige Neuentwicklung eines entscheidungsanalytischen Modells sinnvoll und zeitlich durchführbar ist.

Werden hochwertige internationale entscheidungsanalytische Modelle in der Literatur gefunden, die zu einer Adaptation und Weiterentwicklung für ein nationales Modell geeignet wären, sollten die internationalen Modellautoren kontaktiert und idealerweise eine Kollaboration aufgebaut werden. Es sollte geprüft werden, wie transparent über die Modellstruktur und die relevanten Parameter berichtet wird, ob ein Programmcode direkt übernommen werden kann und wie flexibel die Modellstruktur für eine Anpassung ist.

Ist eine Zusammenarbeit mit den Entwicklern internationaler entscheidungsanalytischer Modelle zustande gekommen, kann eine Übertragung und Programmierung der internationalen Modellstruktur in eine anwenderfreundliche Software stattfinden. Im Idealfall erhält man Kopien

der internationalen Modelle in der jeweiligen Programmiersprache und überträgt diese in die ausgewählte Software. Übertragene Modelle sollten vor einer Abänderung auf fehlerhafte Programmierung geprüft werden.

Die internationale Modellstruktur sollte dann auf die Übertragbarkeit auf den nationalen Kontext geprüft und falls notwendig eine Adaptation der entsprechenden Modellstruktur durchgeführt werden. Änderungen in der Modellstruktur können aus verschiedenen Gründen notwendig sein, wie zum Beispiel aufgrund anderer Klassifikationssysteme des Gesundheitszustands oder aufgrund von Testergebnissen, aufgrund der Standardversorgungspraxis ist eine andere Vorgehensweise notwendig oder die Art der verfügbaren Modelldaten bedingt eine andere Modellstruktur.

3.5.4.1.8 Numerische Implementation oder Adaptation der Modellparameter

Ist die Modellstruktur programmiert, sollte geprüft werden, welche Modellparameter aus dem internationalen Modell übertragbar sind und implementiert werden können, welche Parameter an die nationalen Gegebenheiten angepasst werden müssen sowie welche Parameter neu erhoben werden sollten. Die zu überprüfenden Modellparameter sind weitgehend die Übergangswahrscheinlichkeiten im Krankheitsverlauf, die medizinischen Effektparameter, die Mengen- und Preisparameter der medizinischen Versorgung und eventuell die Lebensqualitätsparameter.

3.5.4.1.9 Klinische und / oder Kosteneffektivitätsanalyse

Nach Abschließen der Modellprogrammierung und einem darauf folgenden ausführlichen „Debugging“-Prozess, erfolgt die eigentliche klinische bzw. Kosteneffektivitätsanalyse für die Basisfallsituation.

3.5.4.1.10 Sensitivitätsanalysen für unsichere Parameter oder Parameter aus Expertenschätzungen

Unsichere Parameter und Parameter, die aus Expertenschätzungen entnommen sind, sollten in Sensitivitätsanalysen hinsichtlich des Einflusses einer Variation dieser Parameter auf das Ergebnis der Basisfallanalyse untersucht werden. Grundsätzlich sollten Einweg-Sensitivitätsanalysen für alle Modellparameter durchgeführt werden, um den Einfluss eines einzelnen Parameters auf das Ergebnis zu untersuchen. Darauf folgend sollten Zweiweg- und Mehrweg-Sensitivitätsanalysen für unsichere Parameter durchgeführt werden, um die Auswirkung einer simultanen Variation mehrerer Parameter auf das Basisfallergebnis zu untersuchen. Bei der Variation der Parameter ist auf die Wahl sinnvoller Spannweiten zu achten.

Erweisen sich unbekannte oder unsichere Parameter bzw. Annahmen als bedeutsam für den Entscheidungsausgang, so ist ein wichtiges Forschungsergebnis die Feststellung des diesbezüglichen Forschungsbedarfs. Sind die Ergebnisse der EA allerdings robust, so kommt der Unsicherheit in Bezug auf diese Parameter oder Annahmen keine limitierende Bedeutung für die untersuchte Fragestellung zu.

3.5.4.1.11 Ergebnisinterpretation und Schlussfolgerungen im Gesamtkontext

Die Ergebnisse der EA sollten im Gesamtkontext entsprechend der Ausgangsfragestellung interpretiert und diskutiert werden. Limitationen und Modellannahmen sollten aufgezeigt und ihr möglicher Einfluss qualitativ diskutiert werden. Ein Vergleich der Ergebnisse mit international publizierter Evidenz ist anzuraten. Aus den Ergebnissen der Analyse sollten für die gegebene Problemstellung Schlussfolgerungen gezogen werden und es sollten nach Übertragung der Evidenz Empfehlungen an die Entscheidungsträger gegeben werden.

3.5.4.1.12 Festlegung des weiteren Forschungsbedarfs

Anhand der entscheidungsrelevanten Parameter, die in Sensitivitätsanalysen oder „Value-of-Information“-Analysen identifiziert wurden, sollte der notwendige zukünftige Forschungsbedarf festgelegt und beschrieben werden.

3.5.4.1.13 Software für Entscheidungsanalysen

Es folgen einige Hinweise für die Verwendung von Softwarepaketen und die programmiertechnische Umsetzung einer Entscheidungsanalyse. Inzwischen existieren komfortable Programmpakete, die die Analyse von Entscheidungsbäumen und Markov-Modellen unterstützen.

Die Erstellung der Struktur des Entscheidungsbaums ist mit Hilfe entscheidungsanalytischer Software sehr anwenderfreundlich und die Erstellung von Ergebnisgraphiken ist Teil des Leistungsprofils dieser Programme. Ein Beispiel ist das Programm DATA Professional (2004, TreeAge Software, Inc., Williamstown, MA, USA), das auch für die Analysen und Graphiken dieses Beitrags verwendet wird. Weitere Programme sind Decision Maker for Windows (1998-2002, S.G. Pauker, F.A. Sonnenberg, J.B. Wong, G. Hagerty, Pratt Medical Group, Boston, MA and University of Medicine and Dentistry of New Jersey, New Brunswick, NJ) und SMLTREE (1989, J. Hollenberg).

Für sehr komplexe Entscheidungsanalysen kann es sinnvoll sein, die Analyse in einer allgemeinen Statistiksoftware (z.B. SAS, SAS Institute Inc., 1999) oder einer Programmiersprache (z.B. C+) zu programmieren. Einfachere Entscheidungsbaumanalysen könnten theoretisch auch in einem Tabellenkalkulationsprogramm (z.B. Microsoft® Excel, Microsoft Corporation) programmiert werden, jedoch besteht hier die große Gefahr, dass aufgrund mangelnder Übersicht bei Programmänderungen nicht alle betroffenen Zellen innerhalb der Kalkulationsblätter oder Arbeitsmappe berücksichtigt werden und diese Fehler unerkannt bleiben.

Ein systematischer Modellvalidierungsprozess und das gründliche Testen auf Programmierfehler (Debugging) ist wesentlicher Bestandteil der Analysearbeit. Dies gilt für alle Softwarepakete.

3.6 Diskussion

(Verfasser: U. Siebert)

3.6.1 Literaturrecherche

In der systematischen Literaturrecherche können nur wenige Beispiele für den Einsatz entscheidungsanalytischer Methoden zum Zweck der Übertragung internationaler Evidenz auf den nationalen Versorgungskontext identifiziert werden. Andererseits propagieren theoretische Artikel und Lehrbücher diese Vorgehensweise.

Dass nur wenige Fallbeispiele für die quantitativ-formale Evidenzübertragung vorliegen, kann drei Gründe haben:

Die Methode befindet sich noch im Entwicklungsstadium. Dafür spricht, dass erst in den letzten Jahren die entscheidungsanalytische Software eine Ebene der Bedienerfreundlichkeit erreicht hat, die einen effizienten Einsatz von Entscheidungsanalysen im Rahmen von HTA und Evidenzbasierter Medizin (EbM) erlaubt.

1. Die Methode ist zu komplex und besitzt zu viele praktische sowie theoretische Limitationen, als dass sie routinemäßig angewendet werden kann. Auf die Limitationen der Methode wird weiter unten in der Diskussion eingegangen (S. Kapitel „Limitationen und

kritische Würdigung des entscheidungsanalytischen Ansatzes“).

2. Es existieren weitere Anwendungsbeispiele, aber die Literaturrecherche ist nicht sensitiv genug, um diese zu identifizieren. Für diese Erklärung spricht, dass das einzige identifizierte entscheidungsanalytische Anwendungsbeispiel zur Evidenzübertragung²⁴ nicht in der systematischen Literaturrecherche sondern über Verweise in Quellenverzeichnissen gefunden wird. Späth et al.⁷¹ stellen fest, dass gesundheitsökonomische Literatur nur zu 73 % in MEDLINE aufgefunden wird. Da sich entscheidungsanalytische Modelle besonders für die Anpassung von Kosten im Rahmen von Kosteneffektivitätsanalysen eignen, wird die Suche auch auf die ökonomischen Datenbank EconLit ausgedehnt, jedoch ohne weitere Treffer zu liefern.

Insgesamt ist zwar nicht auszuschließen, dass weitere Anwendungsbeispiele zur recherchierten Thematik existieren, allerdings ist davon auszugehen, dass es sich um wenige ausgewählte Anwendungen handelt. Eine sehr sensitive Literatursuche für den deutschen Raum ergibt ebenfalls keine Beispiele, so dass zumindest davon auszugehen ist, dass das Verfahren in Deutschland noch kaum bis keine Anwendung fand. Die Diskrepanz zwischen der Literaturlage und den Empfehlungen in den Lehrbüchern zeigt die hohe Relevanz dieses Forschungsgebiets.

3.6.2 Parameterliste und Stufenkonzept

Im Rahmen dieses Methoden-HTA-Berichts kann ein systematisches Instrumentarium zur Übertragung internationaler Evidenz auf den nationalen Versorgungskontext entwickelt werden. Hierzu gehört eine Parameterliste, die bei der Evidenzübertragung zu prüfen ist. Ein Instrument zur Beurteilung der Übertragbarkeit wird innerhalb des gesundheitsökonomischen Panels der German Scientific Working Group Technology Assessment for Health Care bereits früher von Welte et al.⁷⁷ entwickelt und publiziert.

Es wird ein Modell mit den drei Komponenten Entscheidung, Evidenz und entscheidungsanalytisches Modell aufgestellt, von dem ein Stufenkonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext abgeleitet wird. Dieses Stufenkonzept beschreibt in einzelnen konkreten und praktischen Schritten, wie bei der Durchführung der entscheidungsanalytischen Evidenzübertragung vorzugehen ist und welche Optionen dabei bestehen.

Das Stufenkonzept wurde bereits bei zwei deutschen HTA-Projekten im Auftrag des DIMDI erfolgreich eingesetzt^{67, 68}. Diese Benutzung wird weiter unten beschrieben und diskutiert. Der Gebrauch in zukünftigen HTA-Projekten muss zeigen, ob sich der Ansatz allgemein für die Evidenzübertragung eignet bzw. welche Modifikationen und Erweiterungen dafür erforderlich sind.

3.6.3 Übertragung internationaler Evidenz auf den nationalen Kontext

Im Folgenden soll anhand von Beispielen die besprochenen Ergebnisse zur Übertragung internationaler Evidenz auf den nationalen Kontext (Forschungsfrage 1) diskutiert werden.

3.6.3.1 Misoprostol zur Prophylaxe von Magenulzera bei „Nonsteroidal anti-inflammatory drugs“ zur Therapie von Osteoarthritis

Dieses Beispiel ist ausführlich in einem Lehrbuchkapitel von Greiner et al. beschrieben³³, auf das auch bei der folgenden Darstellung zurückgegriffen wird.

Drummond et al.²³ führen einen internationalen Vergleich zur Kosten-Nutzen-Effektivität von Misoprostol zur Prophylaxe von Magenulzera bei schwach wirksamen Analgetika (nonsteroidal anti-inflammatory drugs (NSAIDs)) zur Therapie von Osteoarthritis durch. Zur Berechnung der Kosteneffektivität dieser Technologie wird für die Länder Belgien, Frankreich, Großbritannien und USA eine Entscheidungsbaumanalyse durchgeführt. Die Modellparameter der klinischen

Effektivität für die einzelnen Zweige des Entscheidungsbaums basieren auf den Ergebnissen einer doppelblinden RCT aus den USA. Die Parameter zur Compliance entstammen der Literatur. Die Kosten der ambulanten Behandlung werden mittels Experteninterviews bestimmt und für die Erhebung der Krankenhauskosten werden entweder vorhandene Datenquellen genutzt oder ergänzende Kostenerhebungen durchgeführt.

Die aus den USA stammende Evidenz zur klinischen Wirksamkeit wird in einem entscheidungsanalytischen Modell (Entscheidungsbaum) mit den jeweiligen verfügbaren nationalen Daten zum Ressourcenverbrauch und den Kosten der einzelnen Behandlungsszenarien zusammengefügt. Dabei wird von der Annahme ausgegangen, dass die klinischen Wirksamkeitsparameter länderübergreifende Gültigkeit besitzen, während die Angaben zum Ressourcenverbrauch und zu den Preisen national erhoben werden müssen.

Dieser Ansatz wird zunehmend verwendet. Die Auswertung der Daten erfolgt dabei entweder zentral für jedes Land gesondert oder dezentral in dem jeweiligen Land. Diese Vorgehensweise ist ein Kompromiss zwischen Erhebungsaufwand, nationaler Aussagefähigkeit und internationaler Vergleichbarkeit der Ergebnisse^{32, 33}.

Allerdings ist die verbleibende Arbeit für die nationale Erhebung der Ressourcenverbräuche und Kosten nicht zu unterschätzen. Vielfach beschränken sich Analysten aus Zeit- und Kostengründen auf exemplarische Daten und verwenden stellvertretend für die nationale Zielpopulation Mengengerüste oder aggregierte Kostendaten aus einzelnen, nicht unbedingt generalisierbaren Quellen (z.B. ein Krankenhaus, wenige Praxen, etc.)^{32, 33}. Dies kann zu einem systematischen Fehler (Bias) bezüglich der erhobenen Daten führen.

Greiner et al. interpretieren das geschilderte Vorgehen in folgender Weise³³:

„Entscheidungsanalytische Verfahren sind insbesondere dann einsetzbar, wenn der Wirkzusammenhang für eine veränderte Wahrscheinlichkeit bestimmter Krankheitsereignisse deutlich ist, wenn beispielsweise mittels präventiv eingesetzter Medikamente ein bestimmtes Krankheitsereignis mit bestimmter Wahrscheinlichkeit vermieden werden kann. Diese Vorgehensweise ist weniger angebracht, wenn der Ressourcenverbrauch beim Auftreten eines bestimmten Behandlungsergebnisses durch die evaluierte Technologie verändert wird. Wenn das neue Medikament beispielsweise nicht präventiv wirken würde, wären die im Entscheidungsbaum implementierten Kosten für eine Ulzeration nicht mehr gleich, sondern unterscheiden sich je nach dem, ob Misoprostol verabreicht wurde oder nicht. Zumindest für Partialanalysen können in solchen Fällen entscheidungsanalytische Verfahren trotzdem sinnvoll sein. Wenn beispielsweise ein Teil des Nutzens einer neuen Gesundheitsleistung darin besteht, die Mortalität zu senken, so können die Kostenwirkungen dieses Aspektes mit einem einfachen Entscheidungsbaum transparent gemacht werden. Die Mortalitätsdaten können dazu gegebenenfalls einer ausländischen klinischen Studie entnommen werden, während die Durchschnittskosten der überlebenden Patienten und die Durchschnittskosten der bei der Behandlung verstorbenen Patienten auf nationaler Ebene erhoben werden.“

3.6.3.2 Antivirale Therapie der chronischen Hepatitis C

Zur antiviralen Therapie der chronischen Hepatitis C wird im Rahmen eines HTA-Berichts im Auftrag der DAHTA@DIMDI ein entscheidungsanalytisches Modell zur Evidenzübertragung auf den Kontext des deutschen Gesundheitssystems angewendet^{68, 80}. Ziel ist eine klinische und ökonomische Evaluation von Langzeitkonsequenzen unter verschiedenen therapeutischen

Strategien. Dabei soll eine systematische und transparente Abwägung erreicht werden zwischen dem medizinischen Nutzen, den zusätzlichen Ressourcen für die antivirale Therapie sowie die hierdurch erwarteten Einsparungen bei der Behandlung von Folgezuständen bei Patienten mit chronischer Hepatitis C⁹⁰.

In diesem HTA-Bericht wird zunächst eine systematische Literaturrecherche zur Effektivität sowie Kosteneffektivität von Kombinationstherapien mit Interferon und Ribavirin durchgeführt. Die Literatur wird nach den Vorgaben der German Scientific Working Group in Health Care systematisch selektiert, mittels der Standardchecklisten bewertet und die Ergebnisse in einer qualitativen Synthese dargestellt.

Ferner wird eine systematische EA durchgeführt. Basierend auf Vorarbeiten der Forschergruppe wird ein entscheidungsanalytisches Modell, das German Hepatitis C Model (GEHMO), für den Kontext des deutschen Versorgungssystems entwickelt, um die medizinischen und gesundheitsökonomischen Langzeiteffekte verschiedener antiviraler Therapiestrategien systematisch und explizit zu analysieren⁹¹. GEHMO ist ein entscheidungsanalytisches Markov-Modell und basiert auf einem publizierten und validierten Markov-Modell zur Progression von Hepatitis C^{6, 79}. Die Modellentwicklung und Anpassung an den deutschen Kontext wird von der GEHMO-Gruppe, einem rund 40-köpfigem Expertennetzwerk durchgeführt, in enger Kooperation mit den Entwicklern des amerikanischen Markov-Modells (Arbeitsgruppe Prof. John Wong, Tufts University School of Medicine, Boston, USA).

Die Grundlage der Modellstruktur kann von dem internationalen Ausgangmodell übernommen werden und wird in der Software Data Professional (2001, TreeAge Software, Inc., Williamstown, MA, USA) programmiert. Anschließend werden die für den deutschen Kontext erforderlichen numerischen und strukturellen Änderungen vorgenommen.

GEHMO beinhaltet sowohl deutsche Lebensqualitätsdaten als auch deutsche ökonomische Daten für den Bereich der chronischen Hepatitis C. Die Lebensqualitätsdaten werden in einer Querschnittsstudie mit 428 an chronischer Hepatitis C erkrankten Patienten mit verschiedenen Instrumenten der Lebensqualitätsdatenerhebung durch Hepatologen und Lebensqualitätsexperten erfasst und für die Verwendung in Kosteneffektivitätsstudien durch statistische Rechenverfahren aufbereitet. Die Ressourcenverbräuche werden basierend auf einem „Patienten-Survey“ erhoben und ergänzt durch Expertenschätzungen aus einem Hepatologen-Panel. Die Ressourcenverbräuche werden mit deutschen Preisen bewertet. Alle Kostendaten werden von den Gesundheitsökonomern für die Implementation in einem Entscheidungsmodell aufbereitet. Es zeigt sich, dass sich die deutschen Kostenparameter für die Versorgung der chronischen Hepatitis C im Vergleich zu anderen Ländern teils deutlich unterscheiden^{92, 93}. Daten der Modellparameter zur Erkrankungsprogression werden der Literatur entnommen und von einem epidemiologischen Expertenpanel geprüft. Die Übergangswahrscheinlichkeit für den Zustand Lebertransplantation wird den deutschen Verhältnissen angepasst. Daten zur medizinischen Effektivität der antiviralen Therapien werden im Gegensatz zum Originalmodell von Wong et al., das detaillierte Originaldaten einer internationalen randomisierten kontrollierten Multicenterstudie verwendet, aus einer umfassenden Metaanalyse eines Cochrane-Reviews der Hepato-Biliary Cochrane Group³⁶ entnommen. Die Autoren der Studie werden kontaktiert, um zusätzliche Dateninformationen für eine sinnvolle Implementation der Evidenz in das Markov-Modell zu erhalten. Mortalitätsdaten für das Sterben aufgrund anderer Ursache als Hepatitis C bedingter Leberkomplikationen werden vom Statistischen Bundesamt in Deutschland bezogen.

Eine gesamtgesellschaftliche Perspektive wird eingenommen. Es wird eine Diskontierung von 3 % jährlich vorgenommen. Mit Hilfe der EA können 20-Jahres-Risiken für die Inzidenz von

Hepatitis C-bedingten Lebererkrankungen, Lebenserwartung, qualitätskorrigierter Lebenserwartung, Lebenszeitkosten und das inkrementelle Kosten-Nutzwert-Verhältnis ermittelt werden.

3.6.3.3 Neue Screeningtechnologien in der Zervixkarzinomfrüherkennung

Im Rahmen eines HTA-Berichts im Auftrag der DAHTA@DIMDI wird eine systematische Übersichtsarbeit zur internationalen Evidenz zur Kosteneffektivität neuer Technologien des Zervixkarzinomscreenings verfasst und die Übertragbarkeit der Daten auf das deutsche Gesundheitssystem anhand von Prüflisten der German Scientific Working Group Technology Assessment for Health Care beurteilt⁶⁸. Die evaluierten Screeningstrategien umfassen das traditionelle Papanicolaou-Verfahren, Dünnschichtpräparationsverfahren, automatisierte (computergestützte) Screeningverfahren und ihre Kombination.

Zur Evidenzübertragung wird ein entscheidungsanalytisches Markov-Modell zur Analyse der medizinischen Effektivität und der Kosteneffektivität verschiedener Screeningstrategien für Deutschland entwickelt.

Es werden eine systematische Literaturrecherche und -selektion durchgeführt, die Studienqualität bewertet, die Übertragbarkeit auf das deutsche Gesundheitssystem beurteilt sowie die Ergebnisse systematisch zusammengefasst. Das entscheidungsanalytische Markov-Modell zur klinischen und ökonomischen Evaluation von Langzeit-Outcomes verschiedener Strategien zum Zervixkarzinomscreening wird in Zusammenarbeit mit den Entwicklern zweier internationaler Markov-Modelle zum Zervixkarzinomscreening (Sue Goldie, Harvard Center for Risk Analysis, Harvard University, Boston, MA, USA und Evan Myers Department of Obstetrics and Gynecology, Center for Clinical Health Policy Research-Evidence-Based Practice Center and Division of General Internal Medicine, Duke University, Durham, NC, USA.) entwickelt. Das für den deutschen Kontext entwickelte Markov-Modell stellt ein Hybrid aus den Strukturen beider internationaler Modelle und eigenen auf deutscher medizinischer Praxis basierenden Strukturen dar. Für das Projekt wird ein deutsches Netzwerk bestehend aus Experten der Bereiche Epidemiologie (verschiedene Tumorregister in Deutschland), Gynäkologie, Gesundheitsökonomie, entscheidungstragende Institutionen, Statistik, Ethik und entscheidungsanalytische Modellierung aufgebaut.

Daten der Übergangswahrscheinlichkeiten zwischen den verschiedenen Gesundheitszuständen werden der internationalen Literatur entnommen sowie das Modell basierend auf deutschen epidemiologische Krebsregisterdaten zur Zervixkarzinominzidenz und zur stadienspezifischen Verteilung angepasst und validiert. Für die Erhebung der Kosten der Zervixkarzinomfrüherkennungsverfahren und der Behandlung von Vorstufen sowie definitiven Karzinomen werden deutsche Kostendaten verwendet. Dabei wird die aktuelle Praxis der Diagnose und der Therapie in Deutschland zugrunde gelegt. Eine gesamtgesellschaftliche Perspektive wird eingenommen sowie Kosten und Effekte mit 3 % jährlich diskontiert. Zielgrößen der Analyse sind entdeckte Karzinomfälle, zervixkarzinomspezifische Mortalität, die Lebenserwartung und die Kosteneffektivität der verschiedenen Screeningstrategien.

3.6.4 Standardsituationen für den Einsatz von Entscheidungsanalysen

Im diesem Abschnitt werden Situationen aufgeführt und diskutiert, in denen EA bevorzugt zum Einsatz kommen bzw. notwendig sind (Forschungsfrage 1). Die folgenden Absätze sind gegliedert nach der entsprechenden Situation, bei der entscheidungsanalytische Modelle verwendet werden können. Dabei ist die Funktion der EA entweder die Synthese von Evidenz aus verschiedenen Studien / Datenquellen, die Modifikation / Anpassung von Daten, die Integration verschiedener Endpunkte und die Extrapolation von Studienergebnissen auf andere

Zeithorizonte, Populationen oder Fragestellungen unter bestimmten Annahmen. Die einzelnen erforderlichen Parameter werden dabei meist aus Studien aus verschiedenen Ländern stammen, es kann sich aber auch um Situationen handeln, in denen die EA Methode zur Integration verschiedener Daten aus einem Land eingesetzt werden kann. Diese Diskussion bezieht der Vollständigkeit halber auch Aspekte ein, die nicht im engeren Sinn unter den Begriff „Health Technology Assessment“ fallen, im weitesten Sinn aber bei HTA eine Rolle spielen.

3.6.4.1 Synthese von Kurz- und Langzeitergebnissen

In internationalen Multicenterstudien werden oft intermediäre Endpunkte untersucht. Auch Metaanalysen stellen oft nur eine Synthese der Ergebnisse aus Studien mit kurzem Zeithorizont dar. Von Interesse sind jedoch Langzeitmorbidity und -mortality. Hierzu müssen häufig die Daten aus RCT mit den Ergebnissen von Beobachtungsstudien längerer Zeithorizonte verknüpft werden. Oft wird vorausgesetzt, dass die Wirksamkeit aus den klinischen Studien nicht länderspezifisch ist. Bezüglich der epidemiologischen Langzeitdaten ist es allerdings ratsam, nationale Daten einzubeziehen (z.B. Krebsregister etc.), da diese sehr viel stärker von den landesspezifischen Parametern wie Soziodemographie, Komorbidität und Risikofaktorverteilung abhängen.

Bei Verdacht auf eine genetisch bedingte Therapieansprechrate müssen auch bei der Übertragung der Daten aus den klinischen Studien entsprechende Anpassungen vorgenommen werden, um einen Pharmacogenomics-Bias zu verhindern.

3.6.4.2 Vergleich von Wirksamkeit und Effektivität

Die Ergebnisse zur Wirksamkeit aus internationalen RCT sind nicht immer auf die Routineverhältnisse des deutschen Versorgungssystems übertragbar. Es ist in verschiedenen Szenarien zu prüfen, wie sich die Veränderung einzelner Parameter auf die Gesamt- bzw. Kosteneffektivität eines medizinischen Verfahrens unter deutschen Routinebedingungen auswirkt.

3.6.4.3 Zusammenfassen verschiedener Outcomes

Die EbM liefert häufig aus verschiedenen Studien verlässliche Ergebnisse zu multiplen Outcomes, die sich aber nicht ohne weiteres synthetisieren lassen, z.B. bessere Heilungsraten und gleichzeitig erhöhte Nebenwirkungsraten unter einer bestimmten Therapie. Unter Verwendung des QALY-Konzepts lassen sich diese Endpunkte zusammenfassen. Dabei ist zu prüfen, ob die disaggregierten Parameter (einzelne Outcomes) für das deutsche Versorgungssystem Gültigkeit haben, oder anzupassen bzw. zu substituieren sind. Idealerweise stammen die Nutzwerte (Utilities) aus Deutschland oder einem Land mit ähnlichem Versicherungssystem und kulturellem Kontext.

3.6.4.4 Berücksichtigung von Patientenpräferenzen

Individuelle Patientenpräferenzen können im Rahmen von „Bedside Decision Making“ und „Shared Decision Making“ für die verschiedenen Nutzen und Risiken der möglichen Behandlungsoptionen berücksichtigt werden. Dieses Einsatzgebiet spielt allerdings im Public Health-Bereich eine untergeordnete Rolle, außer das Public Health-Programm sieht den Einsatz eines patientenorientierten Expertensystems als solches vor. Lässt man die Patienten im Rahmen des Entscheidungsprozesses selbst die individuellen Nutzwerte der verschiedenen Gesundheitszustände bestimmen, entfällt das Problem der Übertragbarkeitsbeurteilung der Patientenpräferenzen. Werden die Nutzwerte allerdings mittels multivariater Regressionsverfahren in Abhängigkeit von Prädiktoren berechnet, sollte die Regressionsformel an einer

deutschen Stichprobe validiert worden sein.

3.6.4.5 Einordnung diagnostischer Studien

Die Ergebnisse von Screening- oder Diagnosestudien werden meist in Sensitivität, Spezifität und ROC-Kurven angegeben. Diese Parameter sind mit den zugehörigen klinischen Konsequenzen wie Langzeitmortalität und -morbidity zu verknüpfen. Bei der Übertragung von diesen diagnostischen Gütewerten ist zu prüfen, ob sich die deutsche Zielpopulation bezüglich des „Casemix“ (Schweregradzusammensetzung der Erkrankten) von derjenigen unterscheidet, für die die diagnostische Studie durchgeführt wird. Ist dies der Fall, muss die Sensitivität entsprechend angepasst werden.

Analog sind die im deutschen Gesundheitssystem geltenden Kosten für Screening und diagnostische Tests mit den entsprechenden erwarteten Kosten für die Versorgung erkrankter sowie nicht erkrankter Fälle zusammenzuführen.

Insbesondere sind in diesem Zusammenhang die klinischen und ökonomischen Konsequenzen falsch positiver bzw. falsch negativer Testergebnisse ins entscheidungsanalytische Modell mit einzubeziehen.

3.6.4.6 Studienplanung und „Value-of-Information“-Analyse

Eine Empfehlung eines HTA-Berichts kann sein, eine große randomisierte Studie im deutschen Kontext durchzuführen. Vor einer solchen Empfehlung ist die ethische Rechtfertigung einer solchen Studie zu prüfen. Dabei kann eine Abschätzung der Effekte anhand einer EA mit (vorläufig) ausländischen Daten erfolgen.

Ferner kann eine auf a-priori-Annahmen basierte EA mit ausländischen Daten und den daraus berechneten Erwartungswerten für die Effekte eine fundierte Fallzahlschätzung ermöglichen.

Im Rahmen der Empfehlungen zum weiteren Forschungsbedarfs innerhalb eines HTA-Berichts kann mit Hilfe einer „Value-of-Information“-Analyse eine aus gesamtgesellschaftlicher Sicht adäquate Fallzahl für eine weitere Studie ermittelt werden, d.h. es wird diejenige Fallzahl bestimmt, die notwendig ist, um die Unsicherheit (Konfidenzintervalle) bezüglich der zu untersuchenden Effektivitäts- und Kostenparameter soweit zu reduzieren, dass die Kosten für jeden zusätzlichen in die Studie eingeschlossenen Fall in einem akzeptablen Verhältnis zum durch die Unsicherheitsreduzierung bedingten erwarteten Gewinn stehen.

3.6.4.7 Ersatz für randomisierte Studien

Häufig sind randomisierte Studien aus ethischen, technischen, zeitlichen oder finanziellen Gründen nicht möglich bzw. stehen die Studienergebnisse zu dem Zeitpunkt, zu dem im Rahmen eines HTA-Berichts eine (vorläufige) Entscheidung getroffen bzw. nahe gelegt werden muss, noch nicht zur Verfügung. Hier symbolisieren entscheidungsanalytisch modellierte Szenarien unter Verwendung geprüfter und gegebenenfalls adaptierter ausländischer Daten eine wichtige Unterstützung im Entscheidungsprozess.

3.6.4.8 Reine länderspezifische Übertragung von Studienergebnissen

Wie mehrfach besprochen, lässt sich bei Unterschieden in der Zielpopulation, des klinischen Settings oder nationalen Versorgungsstandards die Effektivität einer Technologie nicht ohne weiteres auf den deutschen Versorgungskontext übertragen. Insbesondere die Übertragung gesundheitsökonomischer Parameter ist in den seltensten Fällen direkt möglich. Bei einer Vielzahl der am Entscheidungsprozess beteiligten Parameter ist davon auszugehen, dass diese im deutschen Kontext anders ausgeprägt sind als in der im Einzelfall vorliegenden Studie⁷⁷.

Während eine Adjustierung von entsprechenden aggregierten Kostenwerten oder klinischen Ergebnissen nur sehr eingeschränkt oder überhaupt nicht möglich ist, können in einer EA alle disaggregierten Parameter einzeln angepasst werden, wo dies erforderlich ist und Daten aus dem Zielkontext vorliegen. Neben der Übertragung von Ergebnissen aus einem bestimmten Land auf ein anderes Land mit unterschiedlichen Versorgungsstandards oder Kostenstrukturen kann dieses Verfahren auch für die Evidenzübertragung von einem auf ein anderes Setting angewendet werden (z.B. ambulante statt teilstationäre Versorgung).

3.6.4.9 Nationale Projektionen

Mit so genannten „Health Policy“ Modellen können für Versorgungsplanung und Politikberatung wichtige Zielgrößen eines Technologieeinsatzes projiziert werden, wie z.B. erwartete Krankenhaustage, notwendige Operationen, erforderliche Organtransplantationen etc. Ein „Health Policy“ Modell simuliert für eine vollständige heterogene Population (z.B. gesamte Population Deutschlands 2002) unter verschiedenen Szenarien die klinischen Ereignisse und die anfallenden Gesamtkosten im Gesundheitssystem über eine zukünftige Zeitperiode (z.B. 2002-2022).

3.6.4.10 Gesundheitsökonomische Evaluation

In gesundheitsökonomischen Studien werden die Kosten der Effektivität gegenübergestellt (z.B. in Kosten pro QALY). Bereits die Integration von Lebenserwartung und Lebensqualität erfordert oft entscheidungsanalytische Markov-Modelle. Ferner erlauben diese die Integration von ökonomischen Konsequenzen einschließlich einer Extrapolation über den Zeithorizont des RCT hinaus.

Da Preise fast nie zwischen Ländern übertragbar sind, kann die EA im einfachsten Fall zur reinen Preissubstitution unter Beibehaltung des Mengengerüsts aus dem Ursprungsland eingesetzt werden. Ein valideres Verfahren wäre die Anpassung des entscheidungsanalytischen Modells bezüglich des Mengen- und Preisgerüsts.

3.6.5 Limitationen und kritische Würdigung des entscheidungsanalytischen Ansatzes

Stellt man sich die Frage, welche Bedeutung entscheidungsanalytischen Modellierungen im Rahmen von HTA zukommt, so ist zu beachten, dass im Rahmen von HTA-Berichten u.a. Effektivitäts- und ökonomische Aspekte einer medizinischen Technologie zu evaluieren sind. Dabei sollen Ergebnisse zu Morbidität, Mortalität, Lebensqualität, Sicherheit und Kosten eines medizinischen Verfahrens aus verschiedenen Datenquellen systematisch zusammengeführt und übergreifend evaluiert werden, um umfassende Bewertungen und Empfehlungen zum Einsatz der untersuchten Technologien ableiten zu können. Diese Zusammenführung der Daten und die Evaluation der Unsicherheit in Sensitivitätsanalysen ist die Aufgabe der Entscheidungsanalyse.

Aus diesem Grund sehen die von der German Scientific Working Group Technology Assessment for Health Care erstellten methodischen Richtlinien den Einsatz der EA zur Berechnung der Kosteneffektivität von Medizintechnologien vor⁴³.

Der vorliegende Methoden-HTA-Bericht ist ein Beitrag zur praktischen Umsetzung der theoretischen entscheidungsanalytischen Konzepte im Rahmen der praktischen Durchführung von HTA-Berichten in Deutschland.

Bei der Übertragung von Studiendaten auf fremde Gesundheitssysteme halten Greiner et al.^{32, 33} den entscheidungsanalytischen Ansatz für das zu bevorzugende Verfahren. Soweit es

der Studienaufbau und die Fragestellung der Studie zulassen, sollte bei der Übertragung der Evidenz zwischen Gesundheitssystemen der entscheidungsanalytische Ansatz die erste Wahl sein. Auf diese Weise können den unterschiedlichen Kostenstrukturen und den institutionellen Gegebenheiten der einzelnen Gesundheitssysteme am besten Rechnung getragen werden. Je detaillierter die Analyse, d.h. je verzweigter der Entscheidungsbaum ist, desto eher ist es möglich, diese Differenzierungen hinreichend vorzunehmen. Allerdings steigen damit auch die Datenanforderungen und gegebenenfalls die Kosten zur Generierung dieser Daten. Der Kostenvorteil, bereits in klinischen Studien erhobene Daten anderer Ländern verwenden zu können, kann somit relativ rasch verloren gehen. Es bleibt in diesem Fall aber immer noch der zusätzliche Vorteil der internationalen Vergleichbarkeit von Resultaten solcher gesundheitsökonomischer Evaluationen, soweit die Studienprotokolle der einzelnen Untersuchungen in den betreffenden Ländern aufeinander abgestimmt sind.

Ebenso wie bei einer klinischen oder epidemiologischen Studie ist auch bei Entscheidungsanalysen a priori die Fragestellung präzise zu formulieren. Ein- und Ausschlusskriterien bezüglich Publikationen oder sonstiger Informationen sind zu Beginn festzulegen. Ferner ist ein detaillierter Zeitplan mit organisatorischem Ablauf notwendig und ein Studienprotokoll, das die einzelnen Arbeitsschritte sowie die qualitativen und die quantitativen Methoden beschreibt. Statistische Analysestrategien, wie sie beispielsweise bei Metaanalysen oder EA eingesetzt werden, sind detailliert festzulegen und eine Abweichung vom Studienprotokoll ist zu begründen sowie im Bericht bzw. der Publikation zu dokumentieren. Nur wenn diese Punkte beachtet werden, kann dem aus der EA resultierenden Ergebnis der entsprechende Stellenwert zugeordnet werden.

In vielen Fällen werden die wissenschaftliche Gemeinschaft oder ein Forscher(team) nicht in der Lage sein, prospektiv Daten zu erheben. Gründe können mangelnde finanzielle oder zeitliche Ressourcen sein, ethische Gründe oder einfach die Tatsache, dass bis zum Vorliegen von Ergebnissen laufender klinischer Studien eine vorläufige Entscheidung zu treffen ist. Sie muss sich dann aufgrund der besten verfügbaren aktuellen Datenlage ergeben. Eine entscheidungsanalytische Modellierung ist dann oft unvermeidbar¹⁵. Aber selbst wenn eine klinische Studie abgeschlossen ist, besitzt sie aufgrund der gegebenen Rahmenbedingungen (Studienpopulation, Versorgungsstruktur) immer nur eine begrenzte Übertragbarkeit. Dies kann durch Entscheidungsanalysen ausgeglichen werden, indem verschiedenartige disaggregierte Parameter für das Zielland modifiziert oder ersetzt werden und dann im entscheidungsanalytischen Modell zusammengeführt und in Sensitivitätsanalysen systematisch variiert werden.

Eine weitere Motivation für eine EA im Rahmen von HTA ist der begrenzte Untersuchungszeitraum einer jeden klinischen Studie. Wenn umfassende Schlussfolgerungen zu ziehen sind, die eine Berücksichtigung eines längeren Zeithorizonts erfordern, sind weitere (Literatur-)Daten hinzuzuziehen und die Ergebnisse über den Studienzeitraum hinaus zu extrapolieren bzw. zu modellieren. Beispiele sind Spätkomplikationen, Rezidive bei Krebserkrankungen und ein krankheitsabhängiges Langzeitmonitoring.

Die besonderen Vorzüge der EA im Rahmen von HTA liegen in der anschaulichen Strukturierung des Entscheidungsproblems, das in klinischen Studien oft nicht sichtbar wird, da nur der primäre Endpunkt für jeden Interventionsarm einer Studie gemessen wird. Insbesondere bei schlechter Datenlage, beispielsweise im Frühstadium des Lebenszyklus einer Technologie kann eine EA zu einer Schätzung der klinischen und ökonomischen Konsequenzen beitragen.

Ein großer Vorteil der EA gegenüber der klinischen Forschung oder der Feldforschung liegt

darin begründet, dass für klinische oder gesundheitspolitische Entscheidungen Daten aus verschiedenen Disziplinen zusammenzutragen sind. So werden für die medizinische und gesundheitsökonomische Evaluation eines Screeningprogramms unter anderem folgende Parameter benötigt: die Prävalenz der Erkrankung, die Inzidenz, die Mortalität, die Stadienverteilung, die Sensitivität und die Spezifität des Screeningtests, die Effektivität sowie die Risiken von Behandlungsmaßnahmen einschließlich Lebenserwartung und Lebensqualität sowie schließlich die assoziierten Akut- und Langzeitkosten. Diese Informationen stammen meist aus Studien in verschiedenen Ländern zu den diversen klinischen, epidemiologischen und ökonomischen Endpunkten. Diese werden im Rahmen der EA angepasst, zusammengeführt und analysiert. Ferner sind häufig verschiedene Konsequenzen wie Lebenserwartung, Heilungsraten, Nebenwirkungen etc. auf eine gemeinsame Skala (z.B. QALY) zu übertragen, um die Vor- und Nachteile eines medizinischen Verfahrens gegenüber einem anderen quantitativ abzuwägen.

Ferner erlaubt eine EA die Gegenüberstellung beliebig vieler Handlungsstrategien, sofern ausreichende und vergleichbare Daten dafür vorhanden sind. Dieser Aspekt wird der Aufgabe der umfassenden Bewertung in HTA gerecht. Ein Problem besteht allerdings oft darin, dass verschiedene Technologien unter verschiedenen Rahmenbedingungen in verschiedenen Ländern bzw. Gesundheitssystemen durchgeführt werden und eine Integration dieser Daten in ein einheitliches Modell nur unter vielen Annahmen oder eben Anpassungsprozessen bzw. neuen Erhebungen zu einzelnen Parametern möglich ist.

Die gesundheitsökonomische EA erlaubt die Problemanalyse aus unterschiedlichen Perspektiven, wie beispielsweise der des Patienten, des Leistungserbringers, der Krankenkassen oder der gesamten Gesellschaft. Auch dies kommt dem Gedanken von HTA entgegen.

Sensitivitätsanalysen können eingesetzt werden, um die Auswirkungen verschiedener Modellannahmen auf die zu treffende Entscheidung systematisch zu evaluieren. Ein großer Wert von Sensitivitätsanalysen liegt darin, diejenigen unsicheren Parameter zu identifizieren, deren Ausprägungen einen wesentlichen Einfluss auf das Ergebnis haben, d.h. für die die Entscheidung sensitiv ist. Die Identifikation dieser „einflussreichen“ Parameter kann erheblich zur Festlegung weiterer Forschungsvorhaben und der sinnvollen Allokation von Forschungsressourcen beitragen.

Allerdings ist zu berücksichtigen, dass die Anwendung entscheidungsanalytischer Modelle insbesondere bei unsachgemäßer Anwendung und Interpretation einige Gefahren birgt. So ist darauf zu achten, dass nicht aufgrund unrealistischer Annahmen komplexe Zusammenhänge zu sehr vereinfacht oder die Ergebnisse in eine bestimmte Richtung verzerrt werden. Hierbei wirkt sich insbesondere die komplexe Methodik der EA negativ aus, da sie bei unzureichender Dokumentation nicht mehr transparent und damit schwer vermittelbar ist. Für Anwender der Ergebnisse kann ein Modell eine Black Box sein, deren Inhalt nicht nachvollzogen werden kann.

Zusammenfassend ist festzustellen, dass die formale EA bei der Evidenzübertragung im Rahmen deutscher HTA-Berichte, für die umfassende deutsche Daten fehlen und nicht von einer direkten Übertragbarkeit der Studienergebnisse aus dem Ausland ausgegangen werden kann, die einzige Möglichkeit zur strukturierten, systematischen und expliziten Entscheidungsfindung darstellt. Wichtig ist jedoch eine verantwortungsvolle Ergebniskommunikation: Die Ergebnisse basieren auf dem aktuellen Wissensstand und ihre Validität hängt von der Richtigkeit der Modellstruktur, der zugrunde gelegten Annahmen sowie der angepassten Parameter ab. Sensitivitätsanalysen erlauben die Beurteilung der Robustheit der Analyseergebnisse bei Veränderung der Modellannahmen.

3.7 Schlussfolgerungen und Empfehlungen

(Verfasser: U. Siebert)

Die folgenden Schlussfolgerungen und Empfehlungen ergeben sich aus den Ergebnissen dieses Methoden-HTA-Berichts:

- Es besteht eine Diskrepanz zwischen den Lehrbüchern, die den entscheidungsanalytischen Ansatz bei der Übertragung von Evidenz zwischen verschiedenen Gesundheitssystemen propagieren und dem Mangel an Anwendungsbeispielen in der Literatur. Diese Lücke verdeutlicht den Forschungsbedarf zu diesem Thema.
- Es werden der Sinn und Nutzen des Einsatzes von EA im Rahmen der Übertragung internationaler Evidenz auf den nationalen Kontext aufgezeigt.
- Es wird ein systematisches Instrumentarium zur Übertragung internationaler Evidenz auf den nationalen Versorgungskontext entwickelt. Dieses Instrumentarium beinhaltet eine Liste der bei der Evidenzübertragung zu prüfenden Parameter.
- Es wird ein Modell mit den drei Komponenten Entscheidung, Evidenz und entscheidungsanalytisches Modell aufgestellt, von dem ein Stufenkonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext abgeleitet wurde. Dieses Stufenkonzept beschreibt in einzelnen konkreten und praktischen Schritten, wie bei der Durchführung der entscheidungsanalytischen Evidenzübertragung vorzugehen ist und welche Optionen es dabei gibt. Verschiedene Situationen, in denen sich der Einsatz entscheidungsanalytischer Modelle anbietet, werden vorgestellt.
- Das Stufenkonzept erweist sich bei zwei deutschen HTA-Berichten („Evaluation antiviraler Therapien bei chronischer Hepatitis C“, „Evaluation neuer Technologien bei der Früherkennung des Zervixkarzinoms“) als praktikabel und erfolgreich.
- Der Einsatz des entwickelten Frameworks in weiteren HTA-Projekten, bei denen sich die Problematik der Übertragung von Evidenz aus ausländischen Studien auf den deutschen Kontext stellt, wird empfohlen. Der für die Entwicklung von entscheidungsanalytischen Modellen erforderliche zeitliche Aufwand ist bei der Ausschreibung von HTA-Berichten zu berücksichtigen.
- Unsichere Parameter sollten in Sensitivitätsanalysen hinsichtlich ihres Einflusses auf das Entscheidungsergebnis untersucht und transparent dargelegt werden. Das Ergebnis der Sensitivitätsanalysen sollte in Empfehlungen zum Forschungsbedarf eingehen.
- Der Einsatz entscheidungsanalytischer Modelle besitzt viele Limitationen. Sie sowie die Annahmen des jeweiligen entscheidungsanalytischen Modells sollten in jedem HTA-Bericht umfassend und transparent dargestellt werden. Die kritische Diskussion der entscheidungsanalytischen Ergebnisse ist notwendig.
- Die Anwendung in zukünftigen HTA-Projekten muss zeigen, ob sich der Ansatz allgemein für die Evidenzübertragung eignet bzw. welche Modifikationen und Erweiterungen dafür erforderlich sind.

Danksagung

Prof. Peter J. Neumann (Harvard Center for Risk Analysis, Harvard School of Public Health, Boston, MA, USA) und Prof. Mark Lipsitch seien (Dept. for Epidemiology, Harvard School of Public Health, Boston, MA, USA) für die fruchtbaren Diskussionen im Rahmen der Erstellung der Parameterliste gedankt. Dr. Silke Siebert (Program on HTA and Decision Sciences, Ludwig-Maximilians-Universität München) sei für die kritische Durchsicht des Manuskripts gedankt.

4 Anhang

4.1 Literatur

1. Anonymous: **Deutsche Empfehlungen zur gesundheitsökonomischen Evaluation - Revidierte Fassung des Hannoveraner Konsensus. Hannoveraner Konsensus Gruppe.** In: Gesundheitsökonomie und Qualitätsmanagement 4 (1999), A62-A65.
2. Baltussen R, Ament A, Leidl R: **Making cost assessments based on RCTs more useful to decision-makers.** In: Health Policy 37 (1996) 3, S. 163-183.
3. Baltussen R, Leidl R, Ament A: **The impact of age on cost-effectiveness ratios and its control in decision making.** In: Health Econ J1 - HE 5 (1996) 3, S. 227-239.
4. Barry MJ, Fleming C, Coley CM, Wasson JH, Fahs MC, Oesterling JE: **Should Medicare Provide Reimbursement for Prostate-Specific Antigen Testing for Early Detection of Prostate Cancer? Part I: Framing the Debate.** In: Urology 46 (1995) 1, S. 2-13.
5. Beck JR, Pauker SG: **The Markov Process in Medical Prognosis.** In: Medical Decision Making 3 (1983), S. 419-458.
6. Bennett WG, Inoue Y, Beck JR, Wong JB, Pauker SG, Davis GL: **Estimates of the cost-effectiveness of a single course of interferon-alpha 2b in patients with histologically mild chronic hepatitis C.** In: Annals of Internal Medicine 127 (1997) 10, S. 855-865.
7. Berlin J: **Benefits of heterogeneity in meta-analysis in data from epidemiologic studies.** In: American Journal of Epidemiology 4 (1995), S. 218-228.
8. Böhmer S, Kohlmann T: **Verfahren zur Bewertung von Gesundheitszuständen und Lebensqualität. In Ravens-Sieberer UC, A: Lebensqualität und Gesundheitsökonomie in der Medizin.** Ecomed, Landsberg, Deutschland, 2000, S. 53-72.
9. Boyle P, Robertson C, Lowe F, Roehrborn C: **Meta-analysis of clinical trials of permixon in the treatment of symptomatic benign prostatic hyperplasia.** In: Urology 55 (2000) 4, S. 533-539.
10. Breyer F, Zweifel P: **Gesundheitsökonomie.** 3. Aufl., Springer, Berlin, Heidelberg, New York, 1999.
11. Briggs A, Goeree R, Blackhouse G, O'Brian BJ: **Probabilistic Analysis of Cost-Effectiveness Models: Choosing between Treatment Strategies for Gastro-esophageal Reflux Disease.** In: Medical Decision Making 22 (2002), S. 290-308.
12. Briggs A, Sculpher M: **An introduction to Markov modelling for economic evaluation.** In: Pharmacoeconomics 13 (1998) 4, S. 397-409.
13. Britton A, Mc Pherson K, Mc Kee M, Sanderson C, Black N, Bain C: **Choosing between randomised and non-randomised studies: a systematic review.** In: Health Technology Assessment (South Hampton, NY) 2 (1998) 13: i-iv, S. 1-124.

14. Buselmaier W, Tariverdian G: **Humangenetik**. 1. Aufl., Springer, Berlin, 1997.
15. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, Vray M: **Modelling in economic evaluation: an unavoidable fact of life**. In: Health Economics 6 (1997) 3, S. 217-227.
16. Colin C, Vergnon P, Jullien AM, Excoffier S, Matillon Y, Trepo C: **Cost-effectiveness of screening blood donors for hepatitis C and non-A, non-B, non-C hepatitis**. In: Journal of Clinical Microbiology and Infectious Diseases 16 (1997) 3, S. 220-227.
17. Corzillius M, Mühlberger N, Sroczynski G, Jaeger H, Wasem J, Siebert U: **Cost effectiveness analysis of routine use of genotypic antiretroviral resistance testing after failure of antiretroviral treatment for HIV**. In: Antiviral Therapy 9 (2004) 1, S. 27-36.
18. Corzillius M, Mühlberger N, Sroczynski G, Peeters J, Siebert U, Jäger H, Wasem J: **Wertigkeit des Einsatzes der genotypischen und phänotypischen HIV-Resistenzbestimmung im Rahmen der Behandlung von HIV-infizierten Patienten. Systematischer Review und entscheidungsanalytische Modellierung zur medizinischen Effektivität und Kosteneffektivität**. Health Technology Assessment Vol 28. Schriftenreihe des Deutschen Instituts für Medizinische Dokumentation und Information im Auftrag des Deutschen Bundesministeriums für Gesundheit und Soziale Sicherung. Asgard, Sankt Augustin, 2002.
19. Leidl RM: **Some factors to consider when using the results of economic evaluation studies at the population level**. In: International Journal Of Health Technology Assessment In Health Care (1994), 10(3), S. 467-478.
20. Dickerson K, Berlin J: **Meta-analysis: state-of-the science**. In: Epidemiologic Reviews 14 (1992), S. 154-176.
21. Djulbegovic B, Hozo I, Lyman GH: **Linking evidence-based medicine therapeutic summary measures to clinical decision analysis 2000**.
22. Doubilet P, Beck CB, Weinstein MC, Braun P, McNeil BJ: **Probabilistic Sensitivity Analysis Using Monte Carlo Simulation. A Practical Approach**. In: Medical Decision Making 5 (1985) 2, S. 157-177.
23. Drummond M, Bloom BS, Carrin G, Hillman AL, Hutchings HC, Knill-Jones RP, de Pouvourville G, Torfs K: **Issues in the cross-national assessment of health technology**. In: Int J Tech Assess Health Care 8 (1992) 4, S. 671-682.
24. Drummond M, O'Brian B, Stoddart GL, Torrance GW: **Methods for the economic evaluation of health care programmes**. 2. Ed., Oxford University Press, Oxford, New York, 1997.
25. Gandjour A, Lauterbach KW: **Zur Wirtschaftlichkeit von Qualitätsverbesserungsmaßnahmen im Gesundheitswesen**. In: Medizinische Klinik 97 (2002) 8, S. 499-502.
26. Gibis B, Busse R, Reese E, Richter K, Schwartz FW, Köbberling J: **Das Mammographie-Screening zur Brustfrüherkennung**. 1998.

27. Gibis B, Rheinberger P: **Erfahrungen mit und Impact von Health Technology Assessment im Bundesausschuss der Ärzte und Krankenkassen.** In: Z ärztl Fortbild 96 (2002), S. 82-90.
28. Gold MR, Siegel JE, Russell LB, Weinstein MC: **Cost-effectiveness in health and medicine.** 1. Ed., Oxford University Press, New York, Oxford, 1996.
29. Graham ID, Stiell IG, Laupacis A, McAuley L, Howell M, Clancy M, Durieux P, Simon N, Emparanza JI, Aginaga JR, O'Connor A, Wells G: **Awareness and use of the Ottawa ankle and knee rules in 5 countries: can publication alone be enough to change practice?** In: Ann Emerg Med 37 (2001) 3, S. 259-266.
30. Greenland S: **Quantitative methods in the review of epidemiologic literature.** In: Epidemiologic Reviews 9 (1987d), S. 1-30.
31. Greenland S: **A critical look at some popular meta-analytic methods.** In: American Journal of Epidemiology 140 (1994d), S. 290-296.
32. Greiner W, Schöffski O, Schulenburg Graf v. d. JM: **The transferability of international economic health-economic results to national study questions.** In: Eur J Health Economics 1 (2000), S. 94-102.
33. Greiner W, Schöffski O, Schulenburg Graf v. d. JM: **Die Übertragbarkeit internationaler Ergebnisse auf nationale Fragestellungen.** In: Schöffski O, Schulenburg Graf v. d. JM: Gesundheitsökonomische Evaluationen. Springer-Verlag, Berlin-Heidelberg, 2000.
34. Haycox A, Dubois D, Butterworth M: **Customising an international disease management model to the needs of individual countries.** In: Pharmacoeconomics 14 (1998) Suppl. 2, S. 39-56.
35. Hertz-Picciotto I, Neutra R: **Resolving discrepancies among studies: the influence of dose on effect size.** In: Epidemiology 5 (1994), S. 156-163.
36. Kjaergard L, Krogsgaard K, Gluud C: **Ribavirin with or without alpha interferon versus no intervention, placebo or alpha interferon for chronic hepatitis C (Cochrane Review).** The Cochrane Library. Oxford, 2000.
37. Kleter B, Brouwer J, Nevens F, van Doorn L, Elewaut A, Versieck J, Michielsen P, Hautekeete M, Chamuleau R, Brenard R, Bourgeois N, Adler M, Quint W, Bronkhorst C, Heijntink R, Hop W, Fevery J, Schalm S: **Hepatitis C virus genotypes: epidemiological and clinical associations. Benelux Study Group on Treatment of Chronic Hepatitis C.** In: Liver 18 (1998) 1, S. 32-38.
38. König HH, Stratmann D, Leidl R: **Effektivität der Kosten medizinischer Leistungen. Grundprinzipien und Qualitätskriterien der ökonomischen Evaluation.** In: Perleth M, Antes G: Evidenz-Basierte Medizin. Wissenschaft im Praxisalltag. MMV, München, 1998, S. 84-93.

39. Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS: **Screening for Prostate Cancer. A Decision Analytic View.** In: J Am Med Assoc 272 (1994) 10, S. 773-780.
40. Kunz R, Oxman A: **The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials.** In: BMJ 317 (1998), S. 1185-1190.
41. Leidl R: **Der Effizienz auf der Spur: Eine Einführung in die ökonomische Evaluation.** In: Schwartz FW et al.: Das Public Health Buch: Gesundheit und Gesundheitswesen. Urban & Schwarzenberg, 1998, 346-369.
42. Light R, Pillemer D: **Summing up: the science of reviewing research.** Harvard University Press, Cambridge, MA., 1984.
43. Lühmann D, Siebert U, Droste S, Von der Leyen U, Perleth M: **Konzeption zur Anpassung und Erstellung von Technologiebewertungen in der deutschen Arbeitsgruppe Verfahrensbewertung in der Medizin.** In: Care GSWGTAfH: Toolkit - Informationsmaterial für Verfasser von HTA-Berichten. German Scientific Working Group Technology Assessment for Health Care, 1999.
44. Mandelblatt J, Fryback D, Weinstein M, Russell L, Gold M, and members of the Panel on Cost-Effectiveness in Health and Medicine: **Assessing the effectiveness of health interventions for cost-effectiveness analysis.** In: Annals of General Internal Medicine 12 (1997) 9, S. 551-558.
45. Mandelblatt J, Fryback D, Weinstein M, Russell L, Gold M, Hadorn D: **Assessing the effectiveness of health interventions.** In: Gold MR, Siegel JE, Russell LB, Weinstein MC: Cost-effectiveness in health and medicine. Oxford University Press, New York, 1996, S. 135-175.
46. Marchioli R, Marfisi RM, Carinci F, Tognoni G: **Meta-analysis, clinical trials, and transferability of research results into practice. The case of cholesterol-lowering interventions in the secondary prevention of coronary heart disease.** In: Archives of Internal Medicine 156 (1996) 11, S. 1158-1172.
47. Mc Kee M, Britton A, Black N, Mc Pherson K, Sanderson C, Bain C: **Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies.** In: BMJ 319 (1999) 7205, S. 312-315.
48. McCrory DC, Matchar DB: **Evaluation of cervical cytology - systematic review.** 1999 Agency for Health Care Research and Quality (AHRQ).
49. Miller DK, Homan SM: **Determining transition probabilities: Confusion and suggestions.** In: Medical Decision Making 14 (1994), S. 52-58.
50. MSD Sharp, Dohme GmbH: **MSD-Manual der Diagnostik und Therapie.** 4. Aufl., Urban & Schwarzenberg, München Wien Baltimore, 1988.

51. Mugford M, Hutton G, Fox-Rushby J: **Methods for economic evaluation alongside a multicentre trial in developing countries: a case study from the WHO Antenatal Care Randomised Controlled Trial.** In: Paediatric & Perinatal Epidemiology 12 (1998) Suppl. 2, S. 75-97.
52. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD: **Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review.** In: Ann Intern Med 6 (2000) 132(10), S. 810-819.
53. Neumann PJ, Hermann RC, Kuntz KM, Araki SS, Duff SB, Leon J, Berenbaum PA, Goldman PA, Williams LW, Weinstein MC: **Cost-effectiveness of donepezil in the treatment of mild or moderate Alzheimer's disease.** In: Neurology 52 (1999), S. 1138-1145.
54. Nugent WC: **Clinical Applications of risk-assessment protocols in the management of individual patients.** In: Ann Thorac Surg 64 (1997) Suppl, S68-S72.
55. Petitti DB: **Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis.** 1st, Oxford University Press, New York-Oxford, 1994.
56. Postma MJ, Kornarou H, Papanizos V, Leidl R, Tolley K, Kyriopoulos J, Jager JC: **Hospital care for persons with AIDS in European-Union countries; a cross-country comparison.** In: Health Care Manag Sci 3 (2000), S. 1-7.
57. Rogers S, Farlow M, Doody R, Mohs R, Friedhoff L, the Donepezil Study Group: **A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease.** In: Neurology 50 (1998), S. 136-145.
58. Rubin D: **A new perspective.** In: Wachter K, Straf M: The future of meta-analysis. Russell Sage Foundation, New York, 1990b, S. 155-165.
59. Rubin D: **Meta-analysis: literature synthesis or effect-size estimation?** In: Journal of Educational Statistics 17 (1992), S. 363-374.
60. Schiffman M, Herrero R, Hildesheim A, Sherman ME, Bratti M, Wacholder S, Alfaro M, Hutchinson M, Morales J, Greenberg MD, Lorinez AT: **HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica.** In: JAMA 283 (2000) 1, S. 87-93.
61. Schöffski O, Schulenburg Graf von der JM: **Gesundheitsökonomische Evaluationen.** 2., Springer, Berlin, 2000.
62. Schwartz F, Badura B, Leidl R, Raspe H, Siegrist J: **Das Public Health Buch.** 1. Aufl., Urban & Fischer, München, Jena, 1998.
63. Schwartz J, Ball J, Moser R: **Safety, efficacy, and effectiveness of clinical practices: a new initiative.** In: Annals of Internal Medicine 96 (1982) 2, S. 246-247.

64. Shani M: **The impact of information on medical thinking and health care policy.** In: Int J Med Inform 58-59 (2000), S. 3-10.
65. Siebert U, Behrend C, Mühlberger N, Wasem J: **Medizinische und gesundheitsökonomische Evidenz des PSA-Screenings beim Prostatakarzinom. Systematischer Review im Rahmen von Health Technology Assessment in Deutschland.** In: Das Gesundheitswesen 61 (1999) A184.
66. Siebert U, Goldie S, Kuntz K: **Pharmacogenomics-Bias. Simulationen zur Fehlbeurteilung von Arzneimitteleffekten bei Nichtberücksichtigung genetischer Heterogenität. [Abstract].** GMDS Tagung, Berlin, (2002).
67. Siebert U, Sroczynski G: **Antivirale Therapie bei Patienten mit chronischer Hepatitis C in Deutschland. Medizinische und ökonomische Evaluation der initialen Kombinationstherapie mit Interferon / Peginterferon und Ribavirin.** Health Technology Assessment Vol 8. Schriftenreihe des Deutschen Instituts für Medizinische Dokumentation und Information im Auftrag des Deutschen Bundesministeriums für Gesundheit und Soziale Sicherung. DIMDI, Köln, 2003.
68. Siebert U, Muth C, Sroczynski G, Velasco-Garrido M, Gerhardus A, Gibis B (Hrsg.) **Dünnschichtpräparationen und computergestützte Untersuchungen von Zervixabstrichen im Rahmen der Krebsfrüherkennung - Medizinische Effektivität, gesundheitsökonomische Evaluation und systematische Entscheidungsanalyse.** Health Technology Assessment Vol 35. Schriftenreihe des Deutschen Instituts für Medizinische Dokumentation und Information im Auftrag des Deutschen Bundesministeriums für Gesundheit und Soziale Sicherung, Vol 35. Asgard, Sankt Augustin, 2004.
69. Sloan FA: **Valuing Health Care, Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies.** Cambridge University Press, Cambridge, 1995.
70. Sonnenberg FA, Beck JR: **Markov models in medical decision making: a practical guide.** In: Medical Decision Making 13 (1993) 4, S. 322-338.
71. Späth HM, Carrre MO, Fervers B, Philip T: **Analysis of the eligibility of published economic evaluations for transfer to a given health care system. Methodological approach and application to the French health care system.** In: Health Policy 49 (1999), S. 161-177.
72. Szucs TD: **Medizinische Ökonomie. Eine Einführung.** Urban & Vogel, München, 1997.
73. Thompson S: **Why sources of heterogeneity in meta-analysis should be investigated.** In: BMJ 309 (1994), S. 1351-1355.

74. Villar J, Ba'aqueel H, Piaggio G, Lumbiganon P, Belinza PM, Farnot U, Al-Mazrou Y, Carroli G, Pinol A, Donner A, Langer A, Nigenda G, Mugford M, Fox-Rushby J, Hutton G, Bergsjö P, Bakketeig L, Berendes H, Group WACTR: **WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care.** In: Lancet 357 (2001) 9268, S. 1551-1564.
75. Wasem J, Siebert U: **Gesundheitsökonomische Parameter einer Evidence-based medicine.** In: Zeitschrift für Ärztliche Fortbildung und Qualitätssicherung 93 (1999) 6, S. 427-436.
76. Weinstein MC, Fineberg HV, Elstein AS, Frazier HS, Neuhauser D, Neutra RR, McNeil BJ: **Clinical Decision Analysis.** 1st, W. B. Saunders Company, Philadelphia, 1980.
77. Welte R, Leidl R: **Übertragung der Ergebnisse ökonomischer Evaluationsstudien aus dem Ausland auf Deutschland: Probleme und Lösungsansätze.** In: Leidl R, Schulenburg Graf v. d. JM, Wasem J: Ansätze und Methoden der ökonomischen Evaluation - eine internationale Perspektive. Nomos, Baden-Baden, 1999, S. 146-168.
78. Willke RJ, Glick HA, Polsky D, Schulman K: **Estimating country-specific cost-effectiveness from multinational clinical trials.** In: Health Econ J1 - HE 7 (1998), S. 481-493.
79. Wong JB, Bennett WG, Koff RS, Pauker SG: **Pre-treatment evaluation of chronic hepatitis C: Risks, benefits and costs.** In: JAMA 280 (1998) 24, S: 2088-2093.
80. Siebert U, Sroczyński G, on behalf of the German Hepatitis C Model (GEHMO) Group and the HTA Expert Panel on Hepatitis C: **Effectiveness and Cost-Effectiveness of Initial Combination Therapy with Interferon / Peginterferon plus Ribavirin in Patients with Chronic Hepatitis C in Germany: A Health Technology Assessment Commissioned by the German Federal Ministry of Health and Social Security.** Int J Tech Assess Health Care (2005), 21(1). S. 55-65.
81. Kunze S, Schnell-Inderst P, Hessel F, Grill E, Nickisch A, Siebert U, Von Voß H, Wasem J, Arbeitsgruppe Health Technology Assessment Neugeborenen-Hörscreening: **Hörscreening für Neugeborene. Health Technology Assessment der medizinischen Effektivität und ökonomischen Effizienz.** Health Technology Assessment Vol 12. Schriftenreihe des Deutschen Instituts für Medizinische Dokumentation und Information im Auftrag des Deutschen Bundesministeriums für Gesundheit und Soziale Sicherung. 2004. Niebüll, Köln.
82. Grill E, Hessel F, Siebert U, Schnell-Inderst P, Kunze S, Nickisch A, Wasem J: **Comparing the clinical effectiveness of different new-born hearing screening strategies. A decision analysis.** In: BMC Pediatrics (2005) [im Druck].
83. Siebert U: **Transparente Entscheidungen in Public Health mittels systematischer Entscheidungsanalyse.** In: Schwartz FW, Badura B, Leidl R, Raspe H, Siegrist J, Walter U, (Hrsg.): Das Public Health Buch. Gesundheit und Gesundheitswesen. 2. Auflage, München: Urban & Fischer (2003), S. 485-502.

84. Siebert U: **When should decision-analytic modelling be used in the economic evaluation of health care? [Editorial]**. In: European Journal of Health Economics (2003), 4(3), S. 143-150.
85. Helou A, Lorenz W, Ollenschläger G, Reinauer H, Schwartz FW: **Methodische Standards der Entwicklung evidenz-basierter Leitlinien in Deutschland. Konsens zwischen Wissenschaft, Selbstverwaltung und Praxis**. In: Zeitschrift für Ärztliche Fortbildung und Qualitätssicherung (2000), 94(5), S. 330-339.
86. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften & Ärztlichen Zentralstelle Qualitätssicherung: **Das Leitlinien-Manual. Entwicklung und Implementierung von Leitlinien in der Medizin**. In: Zeitschrift für Ärztliche Fortbildung und Qualitätssicherung (2001), 95 Suppl. 1, S. 1-84.
87. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften: **Erarbeitung von Leitlinien für Diagnostik und Therapie. Methodische Empfehlungen**. Stand 2000. <http://leitlinien.net> (31.12.2004)
88. Van Damme P, Kane M, Meheus A: **Integration of hepatitis B vaccination into national immunisation programmes**. In: BMJ (1997), 314(7086), S. 1033.
89. Robert Koch-Institut: **Merkblätter für Ärzte: Hepatitis A. Vol. 2001**. 2001.
90. Siebert U: **Public Health-Relevanz und Gesundheitsökonomie der Hepatitis C in Deutschland**. In: Public Health Forum (2002), 10(35), S. 11.
91. German Hepatitis C Model (GEHMO) Group. German Hepatitis C Model (GEHMO). **Klinische und ökonomische Evaluation der antiviralen Therapie bei chronischer Hepatitis C**. In: Forum Public Health (2002), 10(35), S. 12-13.
92. Siebert U, Sroczynski G, Rossol S, Wasem J, Ravens-Sieberer U, Kurth BM, Manns MP, McHutchison JG, Wong JB, German Hepatitis C Model (GEHMO) Group, International Hepatitis Interventional Therapy (IHIT) Group: **Cost effectiveness of peginterferon alpha-2b plus ribavirin versus interferon alpha-2b plus ribavirin for initial treatment of chronic hepatitis C**. In: Gut (2003), 52(3), S. 425-432.
93. Siebert U, Wasem J, Rossol S, Sroczynski G, Aidelburger P, Ravens-Sieberer U, Kurth BM, Manns MP, McHutchison JG, Wong JB: **Antiviral treatment initiation costs in chronic hepatitis C**. In: Gut (2005), 54(1), S. 172-173.

4.2 Abbildungsverzeichnis

Abbildung 1:	Einflussfaktoren beim Entscheidungsprozess.....	9
Abbildung 2:	Stellenwert der Entscheidungsanalyse im Entscheidungsprozess.....	12
Abbildung 3:	Entscheidungsbaum mit Variablen für das Beispiel Test / Behandlung / Keine Behandlung.....	18
Abbildung 4:	Entscheidungsbaum mit Ereignis- und Ein-Jahres-Überlebenswahrscheinlichkeit für das Beispiel Test / Behandlung / Keine Behandlung.....	20
Abbildung 5:	Erwartungswert an einem Knoten.....	21
Abbildung 6:	Ausmitteln und Zurückfalten am Entscheidungsbaum.....	22

Abbildung 7:	Verschiedene generische Entscheidungsbaumtypen zur Modellierung von medizinischer Effektivität unter Unsicherheit.....	24
Abbildung 8:	Entscheidungsbaum mit Ereigniswahrscheinlichkeiten und Kosten für das Beispiel Test / Behandlung / Keine Behandlung.....	26
Abbildung 9:	Ausgemittelter und zurückgefalteter Entscheidungsbaum für die Kosten für das Beispiel Test / Behandlung / Keine Behandlung.....	27
Abbildung 10:	Blasendiagramm zur Veranschaulichung eines Markov-Modells.....	31
Abbildung 11:	Blasendiagramm mit Übergangswahrscheinlichkeiten und Nutzwerten (U1 bis U3) ohne und mit Prävention.	33
Abbildung 12:	Schematische Darstellung eines Markov-Modells für eine chronische Erkrankung mit Gesundheitszuständen, Übergangswegen und Übergangswahrscheinlichkeiten für zwei Zyklen ohne und mit Prävention.	33
Abbildung 13:	Veranschaulichung des Markov-Modells für das Fallbeispiel Prävention mit Angabe der knotenspezifischen Parameter. Software: TreeAge Professional for Health Care (Fa. TreeAge, Williamstown, MA, USA).....	38
Abbildung 14:	Analysiertes Markov-Modell für das Fallbeispiel Prävention mit Angabe der erwarteten QALY je Handlungsalternative für einen Zeitraum von 20 Jahreszyklen. Software: TreeAge Professional for Health Care (Fa. TreeAge, Williamstown, MA, USA).....	39
Abbildung 15:	Einweg-Sensitivitätsanalyse für die Prävalenz im Beispiel.	41
Abbildung 16:	Zweiweg-Sensitivitätsanalyse für Letalität und Prävalenz im Beispiel.	43
Abbildung 17:	Tornadodiagramm für den Einfluss der Variablen Prävalenz, Letalität des diagnostischen Tests, Spezifität und Sensitivität auf die erwartete Ein-Jahres-Überlebenswahrscheinlichkeit im Beispiel.	44
Abbildung 18:	Sensitivitätsanalyse des Kosteneffektivitätsverhältnisses für die Testletalität im Beispiel.....	44
Abbildung 19:	Schematische Darstellung des Vorgehens bei der Evidenzübertragung unter Verwendung der Komponenten Entscheidung, Evidenz und Modell.	63

4.3 Tabellenverzeichnis

Tabelle 1:	Begriffe für Ereigniswahrscheinlichkeiten und Konsequenzen im Entscheidungsbaum.	19
Tabelle 2:	Kosteneffektivität der drei Handlungsalternativen im Beispiel.	28
Tabelle 3:	Matrix der Übergangswahrscheinlichkeiten eines Markov-Modells für eine chronische Erkrankung mit den Gesundheitszuständen „Gesund“, „Krank“ und „Tot“ ohne Prävention.....	32

Tabelle 4:	Matrix der Übergangswahrscheinlichkeiten eines Markov-Modells für eine chronische Erkrankung mit den Gesundheitszuständen „Gesund“, „Krank“ und „Tot“ mit Prävention.	32
Tabelle 5:	Markov-Spur für die Handlungsalternative „Ohne Prävention“ für 20 Zyklen (0-19).	34
Tabelle 6:	Rechercheschritte und –ergebnisse in der Literaturlatenbank MEDLINE.....	46
Tabelle 7:	Rechercheschritte und –ergebnisse in der Literaturlatenbank EconLit.....	47
Tabelle 8:	Rechercheschritte und –ergebnisse in der Literaturlatenbank MEDLINE für den deutschen Kontext.	47
Tabelle 9:	Wichtige Parameter bei der Übertragung von Evidenz zwischen verschiedenen Ländern und Versorgungssystemen.	55
Tabelle 10:	Stufekonzept zur Übertragung von internationaler Evidenz auf den nationalen Kontext.	64

4.4 Abkürzungsverzeichnis

ART	Antiretrovirale Therapiekombination
ASCUS	Atypical Squamous Cells of Undetermined Significance
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
ÄZQ	Ärztlichen Zentralstelle Qualitätssicherung
BMGS	Bundesministerium für Gesundheit und Soziale Sicherung
CEAC	Cost-Effectiveness Acceptability Curves
DAHTA	Deutsche Agentur für Health Technology Assessment
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
DRG	Diagnosis Related Groups
E	Erwartungswert
EbM	Evidenzbasierte Medizin
EA	Entscheidungsanalyse
GEHMO	German Hepatitis C Model
HPV	Humanes Papillomavirus
HTA	Health Technology Assessment

IKEV	Inkrementelles Kosteneffektivitätsverhältnis
LSIL	Low-Grade Squamous Intraepithelial Lesion
NPV	Negativer prädiktiver Wert
LJ	Lebensjahr
NSAIDS	Nonsteroidal Anti-inflammatory Drugs; dt.: schwach wirksame Analgetika
PPV	Positiver prädiktiver Wert
QALY	Quality-Adjusted Life Years, dt.: Qualitätskorrigierte Lebensjahre
RCT	Randomised Clinical Trial, dt.: Randomisierte kontrollierte klinische Studie
WHO	World Health Organisation
1J-Survival	Ein-Jahres-Überlebenswahrscheinlichkeit

Uwe Siebert, M.P.H., M.Sc.

Arzt und Gesundheitswissenschaftler, ist Leiter des Program on HTA and Decision Sciences am Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie der Ludwig-Maximilians-Universität München und Direktor des Cardiovascular Research Program am Institute for Technology Assessment, Massachusetts General Hospital, Harvard Medical School, Boston, USA



Sein methodischer Forschungsschwerpunkt ist die Anwendung entscheidungsorientierter Methoden (Decision Sciences) in den Bereichen Health Technology Assessment, Evidenzbasierte Medizin (EbM), Public Health und Gesundheitsökonomie. Die inhaltlichen Schwerpunkte liegen in den Bereichen kardiovaskuläre Erkrankungen, Infektionskrankheiten, neurologische Erkrankungen und Krebsfrüherkennung, sowie in der Entwicklung neuer quantitativer Methoden zur kausalen Inferenz in Epidemiologie und Entscheidungsanalyse. Er lehrt Entscheidungsanalyse, Epidemiologie und HTA an verschiedenen Universitäten in den USA und Europa.

Gaby Sroczynski, M.P.H.

Master of Public Health, ist Biologin und Gesundheitswissenschaftlerin und derzeit tätig im Program for HTA and Decision Sciences (HTADS) am Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie der Ludwig-Maximilians-Universität München.



Zuvor forschte sie für drei Jahre am Harvard Center for Risk Analysis, Harvard School of Public Health, Boston, USA. Ihre wissenschaftlichen Interessen liegen in den Bereichen Health Technology Assessment und quantitative Methoden in der evidenzbasierten Medizin, wie Metaanalysen, entscheidungsanalytische Modellierungen und Kosten-Effektivitätsanalysen. Ihre derzeitigen Forschungsschwerpunkte sind Zervixkrebsfrüherkennung und antivirale Therapie bei chronischer Hepatitis C. Sie unterrichtet Health Technology Assessment und Entscheidungsanalyse im Rahmen der postgraduellen Ausbildung des Program for HTA and Decision Sciences (HTADS) an der Ludwig-Maximilians-Universität München.

Dr. med. Bernhard Bornschein, M.P.H.

Master of Public Health, ist Arzt und Gesundheitswissenschaftler und derzeit tätig im Program for HTA and Decision Sciences (HTADS) am Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie der Ludwig-Maximilians-Universität München und der Arbeitsgruppe Ökonomie in der Neurologie der Neurologischen Klinik an der Rheinischen Friedrich-Wilhelms-Universität Bonn.



Seine wissenschaftlichen Interessen umfassen die Bereiche Biometrie, Epidemiologie, Entscheidungsanalyse, Gesundheitsökonomie und Health Technology Assessment. Derzeit liegen seine Forschungsschwerpunkte im Bereich der Neurologie und der kardiovaskulären Erkrankungen. Er unterrichtet Entscheidungsanalyse im Rahmen der Postgraduellen Ausbildung des Program for HTA and Decision Sciences (HTADS) an der Ludwig-Maximilians-Universität München.

Health Technology Assessment bei DAHTA

Die systematische Bewertung medizinischer Prozesse und Verfahren, Health Technology Assessment (HTA), ist ein wesentliches Instrument für die gesundheitspolitische Entscheidungsfindung. HTA wird als Werkzeug gesundheitspolitischer Beratung seit Mitte der 90er Jahre in der deutschen Gesundheitspolitik eingesetzt.

Die Integration von HTA als Mittel zur Qualitätssicherung im deutschen Gesundheitswesen führte Ende des Jahres 2000 zur Einrichtung der Deutschen Agentur für HTA des DIMDI (DAHTA@DIMDI). Zu ihren Aufgaben gehören die Entwicklung und die Bereitstellung von Informationssystemen, speziellen Datenbanken und HTA-Berichten.

Im Rahmen der Forschungsförderung durch das DIMDI werden HTA-Berichte erstellt, die nicht nur eine wissenschaftliche Grundlage zu medizinischen Aspekten aus Forschung und klinischem Alltag liefern, sondern auch Stellung nehmen zu relevanten ökonomischen, sozialen, ethischen und juristischen Zusammenhängen. Notwendige Voraussetzungen für die Qualität und Aktualität der Berichte sind anerkannte Standards für die Bewertung wissenschaftlicher Texte.

Vorrang haben Themen, für die gesundheitspolitischer Entscheidungsbedarf besteht. Die Inhalte der HTA-Berichte beziehen sich auf folgende Aspekte:

- Experimentelle Wirksamkeit (efficacy)
- Wirksamkeit unter Alltagsbedingungen (effectiveness)
- Vergleichende Bewertung der Wirksamkeit (comparative effectiveness)
- Gesundheitsökonomische Bewertung (efficiency)
- Soziale, rechtliche und ethische Implikationen

Die Themenfindung und die Prioritätensetzung für die zu erstellenden Arbeiten werden durch ein Kuratorium in einem mehrstufigen Verfahren festgelegt. Die Zusammensetzung des Kuratoriums bildet einen repräsentativen Querschnitt durch die Selbstverwaltung des deutschen Gesundheitssystems, ergänzt durch Verbraucher- bzw. Patientenvertreter sowie Vertreter aus der Industrie als Beobachter. Basis der Themenfindung ist eine Themensammlung, die für eine standardisierte Eingabe von Themenvorschlägen öffentlich zugänglich ist (www.dimdi.de – HTA).

Wie erhält man die Berichte?

Eine elektronische und kostenlose Version des Berichts als pdf-Dokument ist über unsere Internetseite erhältlich (<http://www.dimdi.de> – HTA).

Die Buchversion eines Berichts ist über den Buchhandel oder direkt im DIMDI Webshop für jeweils 20 € zzgl. Versandkosten (Inland: 4 €, Ausland: 6 €) bestellbar (www.dimdi.de – DIMDI Webshop). Eine Ausnahme sind die Arbeiten der Medizinischen Hochschule Hannover, die bei der Nomos Verlagsgesellschaft (Band 1 bis 23) und im Asgard-Verlag (Band 24 bis 35) erschienen sind. Auf den jeweiligen Internetseiten der Verlage finden Sie nähere Informationen sowie Bestellmöglichkeiten (www.nomos.de, www.asgard.de).

Zusätzlich finden Sie in der INAHTA-Datenbank Übersetzungen der Zusammen- sowie der Kurzfassung fremdsprachiger HTA-Berichte.

An Ihren Fragen oder Anregungen zu dem vorliegenden Bericht sind die Autoren sehr interessiert. Wir leiten sie gerne an die Autoren weiter und stellen bei Interesse einen Kontakt her.

Unsere Adresse lautet:

Deutsche Agentur für Health Technology Assessment
des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA@DIMDI)
Waisenhausgasse 36-38 a, 50676 Köln
Tel.: +49 221 4724-1, Fax.: +49 221 472-444, E-Mail: dahta@dimdi.de

Gremien von DAHTA@DIMDI
**Kuratorium Health Technology Assessment (HTA) des Deutschen Instituts für
Medizinische Dokumentation und Information (DIMDI)**

(Stand: Juni 2004)

- | | |
|--|--|
| <ul style="list-style-type: none">• Besetzung des Gemeinsamen Bundesausschusses gemäß § 91 SGB V (G-BA) für ärztliche Angelegenheiten• Besetzung des G-BA für die vertragsärztliche Versorgung• Besetzung des G-BA für die vertragszahnärztliche Versorgung• Besetzung des G-BA für die Krankenhausbehandlung• Bundesausschuss der Ärzte und Krankenkassen• Bundesausschuss der Zahnärzte und Krankenkassen• Bundesärztekammer• Kassenärztliche Bundesvereinigung | <ul style="list-style-type: none">• Kassenzahnärztliche Bundesvereinigung• Deutscher Pflegerat• Deutsche Krankenhausgesellschaft• Spitzenverbände der Krankenkassen• Verband der privaten Krankenversicherung• Bundesarbeitsgemeinschaft der PatientInnenstellen und –initiativen• Arbeitsgemeinschaft der wissenschaftlich medizinischen Fachgesellschaften• Wissenschaftlicher Beirat „Health Technology Assessment“ beim DIMDI |
|--|--|

Ständige Gäste sind ein Vertreter / eine Vertreterin folgender Institutionen:

Bundesverband Medizintechnologie e.V.
Bundesverband der Pharmazeutischen Industrie e.V.
Verband forschender Arzneimittelhersteller e.V.

**Wissenschaftlicher Beirat Health Technology Assessment (HTA) des
Deutschen Instituts für Medizinische Dokumentation und Information (DIMDI)**

(Stand: Juni 2004)

Vorsitzender: Prof. Dr. Christian Ohmann

Wissenschaftlicher Fachberater der Arbeitsgemeinschaft der wissenschaftlich medizinischen Fachgesellschaften (AWMF), Universität Düsseldorf

Prof. Dr. Matthias Schrappe

Wissenschaftlicher Fachberater für Methodik

Dr. Regina Kunz

Wissenschaftliche Vertreterin des Koordinierungsausschuss

Dr. Antonius Helou

Wissenschaftlicher Vertreter des Ausschuss Krankenhaus

Dr. Paul Rheinberger

Wissenschaftlicher Vertreter des Bundesausschuss der Ärzte und Krankenkassen

Petra Corvin

Wissenschaftliche Vertreterin des Bundesausschuss der Zahnärzte und Krankenkassen

Prof. Dr. Walter Lehmacher

Wissenschaftlicher Fachberater für Biometrie / Statistik

Prof. Dr. Stefan Felder

Wissenschaftlicher Fachberater für Gesundheitsökonomie

Prof. Dr. Reinhard Busse

Wissenschaftlicher Fachberater für Gesundheitssystemforschung (Public Health)

Dr. Dittmar Padeken

Wissenschaftlicher Fachberater für Medizintechnik

Prof. Dr. Ulrich Jaehde

Wissenschaftlicher Fachberater für Pharmazie

Prof. Dr. Detlef Heidemann

Wissenschaftlicher Fachberater für die Deutsche Gesellschaft für Zahn-, Mund- und Kieferheilkunde

Bis jetzt als Buch erschienene HTA-Berichte

In der Schriftenreihe von DAHTA@DIMDI sind bisher erschienen:

Bd. 1, 2. Auflage

Stellenwert von Ultraschallverfahren im Rahmen der Osteoporoseversorgung. Früherkennung des Frakturrisikos. 2003.

Bd. 2

Methoden der Metaanalyse von diagnostischen Genauigkeitsstudien. 2003.

Bd. 3

Bewertung unterschiedlicher Hygienekonzepte zur Kontrolle von MRSA (Methicillin-resistente Staphylococcus aureus). 2003.

Bd. 4

Medizinische Effektivität beim Einsatz von GP-IIb / IIIa-Rezeptorantagonisten in der Therapie von akuten Koronarsyndromen. 2003.

Bd. 5

Bewertung der Möglichkeiten und Verfahren zur Aufbereitung medizinischer Einwegprodukte. 2003.

Bd. 6

Gesundheitsökonomische „Kurz-HTA-Berichte“. Eine systematische Übersichtsarbeit zur Methodik und Implementation. 2003.

Bd. 7

Glycoprotein IIb / IIIa-Rezeptorantagonisten in der Therapie akuter koronarer Syndrome. Ein gesundheitsökonomischer HTA-Bericht. 2003.

Bd. 8

Antivirale Therapie bei Patienten mit chronischer Hepatitis C in Deutschland. Medizinische und ökonomische Evaluation der initialen Kombinationstherapie mit Interferon / Peginterferon und Ribavirin. 2003.

Bd. 9

Methoden zur Erfassung ethischer Aspekte und gesellschaftlicher Wertvorstellungen in Kurz-HTA-Berichten. Eine internationale Bestandsaufnahme. 2003.

Bd. 10

Ultraschall in der Schwangerschaft. Beurteilung der routinemäßigen Schwangerschafts-ultraschalluntersuchungen unter Maßgabe der Mutterschaftsrichtlinien. 2003.

Bd. 11

Stroke Units: Update des HTA Berichts „Die Evaluation von Stroke Units als medizinische Technologie“. 2004.

Supplement 1

4. Symposium Health Technology Assessment – Bewertung medizinischer Verfahren, 13.-14. November 2003, Krefeld. 2004.

Bd. 12

Hörscreening für Neugeborene. Ein Health Technology Assessment der medizinischen Effektivität und der ökonomischen Effizienz. 2004.

Bd. 13

Wertigkeit von Tumornekrosefaktor-alpha-Antagonisten in der Behandlung der rheumatoiden Arthritis. 2004.

Bd. 16

Der Einfluss von HTA-Berichten auf die gesundheitsökonomische Entscheidungsfindung. 2004.

Die folgenden Berichte sind in der Schriftenreihe HTA des DIMDI erschienen und im Rahmen des Projekts „Aufbau einer Datenbasis für die Evaluation von medizinischen Verfahren und Technologien“ mit der Medizinischen Hochschule Hannover entstanden.

Bd. 1: Bestandsaufnahme, Bewertung und Vorbereitung der Implementation einer Datensammlung. 1998.

Bd. 2: Die Evaluation von Rückenschulprogrammen als medizinische Technologie. 1998.

Bd. 3: Das Mammographiescreening als Verfahren zur Brustkrebsfrüherkennung. 1998.

Bd. 4: Biochemisches Screening für fetale Chromosomenanomalien und Neuralrohrdefekte. Eine Verfahrensbewertung. 2001.

Bd. 5: PSA-Screening beim Prostatakarzinom. 1999.

Bd. 6: Verfahrensbewertung der Magnet-Resonanz-Tomographie (MRT) in der Diagnostik des Mammakarzinoms. 1999.

Bd. 7: Minimal-invasive Therapie der benignen Prostatahyperplasie (BPH-Syndrom). 1999.

Bd. 8: Evaluation präoperativer Routinediagnostik (Röntgenthorax, EKG, Labor) vor elektiven Eingriffen bei Erwachsenen. 1999.

Bd. 9: Stellenwert der Magnet-Resonanz-Tomographie im Rahmen der Versorgung von Patienten mit Rückenschmerzen. Kurz-HTA: Update einer Best-Evidence-Synthese. 1999.

Bd. 10: Stenting versus Ballondilatation bei koronarer Herzkrankheit. 1999.

Bd. 11: Bewertung von Verfahren zur Diagnostik der akuten Sinusitis maxillaris bei Erwachsenen. 1999.

Bd. 12: Spezifische Hyposensibilisierung mit Allergenextrakten bei extrinsischem Asthma bronchiale und Insektengiftallergie. 1999.

Bd. 13: Die Rolle der Osteodensitometrie im Rahmen der Primär-, Sekundär- und Tertiärprävention / Therapie der Osteoporose. 2000.

Bd. 14: Evaluation arthroskopischer Operationen bei akuten und degenerativen Meniskusläsionen. 2000.

Bd. 15: Die Evaluation von Stroke Units als medizinische Technologie. 2000.

Bd. 16: Hochdosis-Chemotherapie mit autologer Stammzelltransplantation zur Therapie des metastasierenden Mammakarzinoms. 2000.

Bd. 17: Vergleichende Effektivität und Differentialindikation von Ballondilatation (PTCA) versus Bypasschirurgie bei Ein- und Mehrgefäßerkrankungen der Herzkranzgefäße. 2000.

Bd. 18: Hüftgelenkendoprothetik bei Osteoarthrose. Eine Verfahrensbewertung. 2000.

Bd. 19: PSA-Screening beim Prostatakarzinom - Systematischer gesundheitsökonomischer Review. Entwicklung und Anwendung eines Instrumentariums zur systematischen Beschreibung und Bewertung gesundheitsökonomischer Studien. 2001.

Bd. 20: Ökonomische Evaluationen der Positronen-Emissions-Tomographie. Ein gesundheitsökonomischer HTA-Bericht. 2001.

Bd. 21: Die Wertigkeit der Stressechokardiographie in der Primärdiagnostik der koronaren Herzkrankheit. 2003.

Bd. 22: Operative Eingriffe an der lumbalen Wirbelsäule bei Bandscheibenbedingten Rücken- und Beinschmerzen. Eine Verfahrensbewertung. 2003.

Bd. 23: Die Verwendung des Excimer Lasers in der refraktiven Augenchirurgie. 2003.

Bd. 24: Positronen-Emissions-Tomographie. Systematische Übersichten zur Wirksam-

keit bei ausgewählten Indikationen. 2003.

Bd. 25: Das Schlafapnoe-Syndrom. Eine systematische Übersichtsarbeit zur Diagnostik, Therapie und Kosteneffizienz. 2003.

Bd. 26: Wertigkeit des Einsatzes der intravaskulären Ultraschallbildgebung (IVUS) im Rahmen von diagnostischen und therapeutischen Herzkatheteruntersuchungen. Medizinische Effektivität. 2003.

Bd. 27: Wertigkeit des Einsatzes der intravaskulären Ultraschallbildgebung (IVUS) im Rahmen von diagnostischen und therapeutischen Herzkatheteruntersuchungen. Ein gesundheitsökonomischer HTA-Bericht. 2003.

Bd. 28: Wertigkeit des Einsatzes der genotypischen und phänotypischen Resistenzbestimmung im Rahmen der Behandlung von HIV-infizierten Patienten. Systematisches Review und entscheidungsanalytische Modellierung zur medizinischen Effektivität und Kosteneffektivität. 2003.

Bd. 29: Verfahrensbewertung der CT-Angiographie, MR-Angiographie, Doppler-Sonographie und Szintigraphie bei der Diagnose von Nierenarterienstenosen. 2003.

Bd. 30: Vergleich verschiedener chirurgischer Verfahren zur elektiven Leistenhernienoperation bei Erwachsenen. Ein Health Technology Assessment. 2003.

Bd. 31: Stenting versus Ballondilatation bei

koronarer Herzkrankheit. Systematische Übersicht zur medizinischen Effektivität und gesundheitsökonomische Bewertung. 2003.

Bd. 32: Knochen- und Knochenersatzmaterialien zur paradontalen Regeneration und zum Knochenaufbau für Implantate. Eine systematische Bewertung der medizinischen Wirksamkeit. 2003.

Bd. 33: Knochen- und Knochenersatzmaterialien zur paradontalen Regeneration und zum Knochenaufbau für Implantate. Ein gesundheitsökonomisches HTA. 2003.

Bd. 34: Methodenmanual für „HTA-Schnellverfahren“ und Exemplarisches „Kurz-HTA“: Die Rolle der quantitativen Ultraschallverfahren bei der Ermittlung des Risikos für osteoporotische Frakturen. 2003.

Bd. 35: Dünnschichtpräparationen und computergestützte Untersuchungen von Zervixabstrichen im Rahmen der Krebsfrüherkennung. Medizinische Effektivität gesundheitsökonomische Evaluation und systematische Entscheidungsanalyse. 2003

Die systematische Bewertung medizinischer Prozesse und Verfahren, *Health Technology Assessment* (HTA), ist mittlerweile integrierter Bestandteil der Gesundheitspolitik. HTA hat sich als wirksames Mittel zur Sicherung der Qualität und Wirtschaftlichkeit im deutschen Gesundheitswesen etabliert.

Seit Einrichtung der Deutschen Agentur für HTA des DIMDI (DAHTA@DIMDI) im Jahr 2000 gehören die Entwicklung und Bereitstellung von Informationssystemen, speziellen Datenbanken und HTA-Berichten zu den Aufgaben des DIMDI.

Im Rahmen der Forschungsförderung beauftragt das DIMDI qualifizierte Wissenschaftler mit der Erstellung von HTA-Berichten, die Aussagen machen zu Nutzen, Risiko, Kosten und Auswirkungen medizinischer Verfahren und Technologien mit Bezug zur gesundheitlichen Versorgung der Bevölkerung. Dabei fallen unter den Begriff Technologie sowohl Medikamente als auch Instrumente, Geräte, Prozeduren, Verfahren sowie Organisationsstrukturen. Vorrang haben dabei Themen, für die gesundheitspolitischer Entscheidungsbedarf besteht.