

Innovationen und ihre Bewertung – eine Frage der Perspektive

von Monika Lelgemann¹

ABSTRACT

Auch in Zeiten von evidenzbasierter Medizin und Health Technology Assessment, also vermeintlich rigorosen methodischen Vorgehensweisen, spielen Grundeinstellungen gegenüber Neuerungen eine entscheidende Rolle in deren Rezeption und Bewertung. Am Beispiel des inzwischen vom Markt genommenen Appetitzüglers Rimonabant werden einige Aspekte beleuchtet, die den Umgang mit Innovationen betreffen. Das Beispiel dient dazu, auf Unzulänglichkeiten hinzuweisen, die gerade bei Modellierungen bestehen, die auf Annahmen oder unsicheren Daten beruhen. Mit dem Konzept des „acceptable regrets“ wird eine Möglichkeit dargestellt, wie mit der bei neuen Technologien immer bestehenden Unsicherheit aufgrund unzulänglicher Datenbasis umgegangen werden kann. Die Kumulation zu vieler Innovationen (zum Beispiel Wirkstoff, Modellierung, Vertragsform) kann zu einer echten Gefahr führen. Dies gilt insbesondere dann, wenn Sorgfältigkeitsanforderungen verletzt und die unsichere Datenbasis nicht transparent dargelegt werden.

Schlüsselwörter: Health Technology Assessment, Rimonabant, Modellierung, evidenzbasierte Medizin, Innovationen

Even in the days of supposedly rigorous methodological approaches such as evidence-based medicine and health technology assessment, basic attitudes towards innovations play an important role in their reception and evaluation. Illustrated by the example of the appetite suppressant Rimonabant, which has been withdrawn from the market, the article highlights some aspects concerning the appraisal and assessment of innovations. The example reveals deficiencies in current models which are based on assumptions or uncertain data. The concept of “acceptable regrets” illustrates a way of handling uncertainties which are always involved with new technologies due to inadequate data. The accumulation of too many innovations (e.g. drug, modelling, contractual form) can be a real hazard. This is especially true if accuracy requirements are violated and the unreliability of the data is not made transparent.

Keywords: Health technology assessment, Rimonabant, modelling, evidence-based medicine, innovations

1 Einleitung

In diesem Beitrag geht es weniger um die Methodik der Innovationsbewertung als vielmehr um die ausgesprochenen und unausgesprochenen Erwartungen, die sich an Innovationen knüpfen. Anhand des Beispiels um die Markteinführung des Appetitzüglers Rimonabant (siehe Seite 17) werden einige Aspekte angesprochen, die beim Umgang mit Innovationen von Bedeutung sind. Die Betrachtung folgt dabei streng unsystematischen Kriterien und erhebt in keiner Weise den Anspruch auf Vollständigkeit. Für systematische Übersichten sei auf andere Autoren und Quellen verwiesen, unter anderem auf die Methodenarbeit im Rahmen des HTA-Programms des

Deutschen Instituts für medizinische Dokumentation und Information (DIMDI). Eine Übersicht (HTA-Bericht), die die systematische Aufarbeitung von Methoden zur frühzeitigen begleitenden Bewertung von Innovationen zum Inhalt hat, wird derzeit erstellt (Hagen und Gerhardus 2008).

2 Health Technology Assessment und Innovationen

Die Entstehung des Health Technology Assessment (HTA) ist eng mit Fragestellungen rund um die Einführung neuer

¹ Monika Lelgemann, MSc, Interdisziplinäres HTA-Zentrum Universität Bremen, c/o Institut für Gesundheits- und Medizinrecht · Postfach 330440 · 28334 Bremen · Telefon: 0421 218-3784 · Telefax: 0421 218-439 · E-Mail: lelgemann@uni-bremen.de

Technologien assoziiert. Neue Technologien werden im Sinne des HTA weit gefasst, sodass darunter sowohl neue Arzneimittel, Medizinprodukte, diagnostische Verfahren, Abrechnungsvorgaben, Anforderungen an medizinische Fortbildungen, Bewertungsverfahren, gesetzliche Regelungen oder – wie im Beispiel Rimonabant (siehe Seite 17) beschrieben – neue Vertragsformen verstanden werden.

Im engeren Kontext fällt auf, dass im angloamerikanischen Sprachraum eine vorsichtigeren Wortwahl üblich ist. Hier wird überwiegend von „new emerging technologies“ und nicht von Innovationen gesprochen. Hinweise auf den stark wertenden Aspekt des Begriffs „Innovation“ gibt die im deutschen Sprachraum traditionsreiche Debatte um Innovationen im Kontext von Arzneimitteln, in der es unter anderem darum geht, zwischen Innovationen, Sprunginnovationen und Scheininnovationen zu differenzieren.

Die enge Verknüpfung von HTA mit dem Thema Innovationen ist multikausal, hat aber nicht zuletzt auch pragmatische Gründe. Da es aufgrund begrenzter Ressourcen unumgänglich ist, eine Priorisierung zu bewertender Technologien durchzuführen, erfolgt eine Fokussierung auf neue Technologien. Diese sind ganz in Analogie zu Arzneimitteln anfänglich in aller Regel mit höheren Kosten und mit Unsicherheit hinsichtlich potenziellem Nutzen und Schaden behaftet. Der Bedarf an einer systematischen Aufarbeitung und Darlegung des zur Verfügung stehenden Wissens, um eine informierte Entscheidung treffen zu können, ist aufgrund dieser Faktoren bei Anwendern, Patienten, Kostenträgern und politischen Entscheidungsträgern besonders groß.

3 Entscheidungen bei unsicherer Datenlage

Das unumgängliche Dilemma besteht darin, dass genau zu dem Zeitpunkt des größten Bedarfs an Entscheidungsunterstützung die Datenlage (body of evidence) begrenzt ist und im Einzelfall immer das Urteil darüber zu fällen ist, ob diese ausreicht, eine begründete Entscheidung für oder gegen eine neue Technologie treffen zu können. Diese Entscheidung wird wesentlich bestimmt sein durch die jeweilig eingenommene Perspektive. Unmittelbar einleuchtend ist der Unterschied der Sichtweise von Anbietern einer neuen Technologie, zum Beispiel den Herstellern, und den Kostenträgern, zum Beispiel den Krankenkassen. Zwar verfolgen sowohl Hersteller als auch Krankenkassen proklamatorisch das gemeinsame Ziel einer Verbesserung der Patientenversorgung, die jeweilige Interessenlage bedingt aber Unterschiede in der Bewertung dessen, was beziehungsweise wie viel als Verbesserung der Versorgung verstanden wird. Zwar ist inzwischen anerkannt, dass bestimmte methodische Vorge-

hensweisen, die den Prinzipien der evidenzbasierten Medizin (EbM) entsprechen, bei der Beurteilung einer Innovation anzuwenden sind. Sie lösen aber die Wertungsfragen, die sich an vielen Stellen ergeben, nicht. Strech spricht von der „impliziten Normativität in medizinischen Nutzevaluierungen“. Nur wenn sich die Anwender der EbM-Methode der Tatsache differierender Werturteile bei unterschiedlichen zugrunde liegenden Interessen bewusst sind, kann eine kritische Diskussion der Konsequenzen erfolgen (Strech 2007).

Eine kontroverse Diskussion besteht auch bei der Frage, wann ausreichend Daten vorliegen, um eine Innovation in die Versorgung einführen zu können. Hier handelt es sich vermutlich vor allem um eine Auseinandersetzung zwischen Skeptikern und Enthusiasten, wobei sich die Skeptiker zufällig häufig im Lager der evidenzbasierten Medizin finden (Donner-Banzhoff 2007). Unterstellt wird, dass bei den Enthusiasten häufig ein Interessenskonflikt vorliegt, der zum Beispiel darin besteht, dass sie an der Entwicklung einer neuen Technologie beteiligt waren. Hiermit ist nicht ausschließlich der Fall finanzieller Interessenkonflikte gemeint, sondern der einer Überzeugung, dass eine bestimmte Maßnahme, beispielsweise eine neue Operationstechnik oder diagnostische Methode, eindeutig besser und den herkömmlichen Verfahren überlegen ist.

Im Rahmen der klinischen Forschung wird in diesem Zusammenhang vom „expectation bias“ oder „optimism bias“ gesprochen, also dem unbegründeten Glauben an die Überlegenheit einer neuen Intervention (Chalmers und Matthews 2006). Aus älteren empirischen Arbeiten ist bekannt, dass dieser Glauben häufiger in herstellerfinanzierten Studien zu finden ist als in öffentlich finanzierten Studien. Wären die positiven Erwartungen an die neuen Interventionen wirklich berechtigt, würden sich in letzter Konsequenz randomisierte Studien verbieten, da sie ethisch nicht vertretbar wären. Den Patienten, denen durch die Randomisierung die neue Intervention vorenthalten wird, würde unvermeidbarer Schaden entstehen. Gleiches gilt für die Annahme der Unbedenklichkeit. Nur wenn diese besteht, ist die randomisierte Prüfung ethisch gerechtfertigt. Das Phänomen „optimism bias“ löst also die entscheidende Frage aus, wie die A-priori-Wahrscheinlichkeit einzuschätzen ist, dass eine zu untersuchende Innovation besser ist als die herkömmliche Intervention. Bisher zum Thema vorliegende Ergebnisse zeigen, dass die Wahrscheinlichkeit der Überlegenheit oder Unterlegenheit der zu testenden neuen Intervention im Rahmen einer randomisierten kontrollierten Studie etwa 50 Prozent beträgt (Djulgovic et al. 2008; Kumar et al. 2005). Eine optimistische Grundhaltung gegenüber Innovationen ist im Zusammenhang mit einer systematischen Evaluation nicht erwünscht, da sie die Gefahr einer systematischen Verzerrung insbesondere der Wahrnehmung von Ergebnissen birgt. Anders ist möglicherweise die Situation

FALLBEISPIEL „RIMONABANT“

Studienlage: Die Markteinführung zahlreicher Appetitzügler ist an Sicherheitsproblemen gescheitert; die aktuell zugelassenen Arzneistoffe Sibutramin und Orlistat haben relevante Verträglichkeitsprobleme. Mit dem Endocannabinoidrezeptor-Antagonisten Rimonabant (Acomplia®) keimte im Jahr 1994 neue Hoffnung auf. Nach viel versprechenden Phase-I- und Phase-II-Studien legte der Hersteller für die Indikation Adipositas die RIO-Studie (RIO = Rimonabant in obesity) mit vier Unterstudien auf. An den Studien nahmen über 4.000 größtenteils weiße Frauen teil, die einen Body-Mass-Index (BMI) > 30 oder einen BMI > 27 und eine Fettstoffwechselstörung haben mussten. Rimonabant 20 und 5 mg wurden doppelblind, randomisiert gegen Placebo untersucht. Das heißt, die Studienteilnehmer wurden nach dem Zufallsprinzip auf die Gruppen verteilt und weder sie noch der Arzt wussten, welcher Gruppe sie zugeordnet waren. Entgegen üblicher Adipositastherapien wurde den Teilnehmern nur eine Kalorienreduktion um 600 Kcal pro Tag empfohlen. Weitere Lebensstilmodifikationen waren nicht verlangt.

Unter Placebo verloren die Probanden im ersten Jahr 2 kg, unter Rimonabant 20 mg 4–5 kg. Nach Umstellung auf Placebo stieg das Körpergewicht rasch wieder an. Langfristige klinische Endpunkte wurden nicht untersucht. Gut 40 % der Probanden brachen die Therapie schon im ersten Jahr ab, aufgrund unerwünschter Arzneimittelwirkungen (UAW) doppelt so viele in der Rimonabant- wie in der Placebo-Gruppe. Das Gewicht der Abbrecher wurde nicht weiter erfasst, sondern mit der LOCF-Methode (LOCF = last observation carried forward) berechnet, wodurch es vermutlich zu einer Überschätzung des Therapieeffekts sowie einer Unterschätzung der UAW-Rate kam. Hinsichtlich unerwünschter Wirkungen fallen unter Rimonabant psychische Störwirkungen auf. Depressionen, Suizidalität, Angst, Schlafstörungen, Krampfanfälle, Stimmungsveränderungen, Reizbarkeit und andere treten unter Rimonabant auf (Placebo: 14 %). Patienten mit Depression als Vorerkrankung waren ausgeschlossen (Christensen et al. 2007).

Rimonabant wurde im Juni 2006 von der Europäischen Arzneimittelagentur (EMA) zugelassen und im September 2006 in Deutschland eingeführt. Im Folgemonat wurde es vom Gemeinsamen Bundesausschuss (GBA) als Lifestyle-Medikament klassifiziert und damit in Deutschland von der Erstattung ausgeschlossen. Die Berater der amerikanischen Arzneimittelzulassungsbehörde FDA kamen aufgrund von Sicherheitsbedenken im Juni 2007 einstimmig zu einer negativen Beurteilung. Die EMA verfügte zweimal (Juli 2007 und Mai 2008) Änderungen der Fachinformation. Dabei wurde unter anderem Depression als Kontraindikation eingeführt und durch Fettdruck hervorgehoben.

Aufbereitete Evidenz: Schon im Oktober 2003 wurde Rimonabant mit zwei anderen Substanzen in einem Cochrane-Review verglichen (Padwal et al. 2004). UAW werden dort als „mood disorders“ summiert, die hohen Abbruchraten problematisiert. Im

Oktober 2006 erscheint ein weiteres Cochrane-Review (Curioni und Andre 2006), das Rimonabant nur in der Indikation Adipositas untersucht. Es sind die gleichen Studien wie im ersten Review enthalten, das Fazit ist vorsichtig positiv. Über UAW wird nur summarisch berichtet, die hohe Abbrechrate nicht thematisiert.

Modellierung: Im Jahr 2006 stellten Aidelsburger und Wasem einen „innovativen Ansatz zur Bestimmung der Kosteneffektivität und der Budgetauswirkung neuer Wirkstoffe am Beispiel Rimonabant“ vor (Aidelsburger und Wasem 2006). Entwickelt wurde ein sogenanntes Health-Policy-Modell, mit welchem sowohl die Kosteneffektivität als auch die Kostenauswirkungen bezogen auf die Gesamtbevölkerung ermittelt werden sollten. Um für das Health-Policy-Modell auf die Gesamtbevölkerung schließen zu können, wurden in einer Querschnittstudie Daten aus über 1.500 Praxen mit 35.000 Patienten erhoben. Die Autoren schlussfolgern, dass auf diese Weise bereits vor Zulassung eines neuen Arzneimittels Ergebnisse zur Kosteneffektivität und zur Budgetwirksamkeit vorliegen können. Inwieweit Wirksamkeit und UAW in die Modelle eingehen kann der Publikation nicht entnommen werden.

Neue Vertragsformen: Unter anderem aufgrund der Ergebnisse der gesundheitsökonomischen Berechnungen wurde ein Kooperationsvertrag zwischen der Deutschen BKK und dem Hersteller geschlossen, der Patienten finanzielle Anreize für das Erreichen von Therapiezielen unter Einnahme von Rimonabant gab (arznei-telegramm 2008). Die Patienten erhielten 30 Euro für das Erreichen des Therapieziels nach drei Monaten und weitere 60 Euro nach insgesamt neun Monaten für das Halten des Therapieziels, wobei nicht vergessen werden darf, dass die Patienten das Arzneimittel aufgrund der GBA-Entscheidung selbst bezahlen müssen.

Marktrücknahme: Am 23. Oktober 2008 veröffentlicht die EMA die Empfehlung, die Zulassung ruhen zu lassen, da die Schädwirkungen häufiger und schwerer und der Nutzen geringer als ursprünglich angenommen seien. Daraufhin nimmt der Hersteller das Präparat vom Markt. Andere Hersteller stellen ihre Forschungsprogramme an anderen Endocannabinoidrezeptor-Wirkstoffen ein.

Literatur

- Aidelsburger P, Wasem J (2006): Ein innovativer Ansatz zur Bestimmung der Kosten-Effektivität und der Budgetauswirkung neuer Wirkstoffe am Beispiel von Rimonabant. 51. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie. Düsseldorf, Köln: German Medical Science (Ref. Type: Conference Proceedings)
- arznei-telegramm (2008): Deutsche BKK und das Lifestyle-Mittel Rimonabant (Acomplia). Jg. 39, 49–50
- Christensen R, Kristensen PK, Bartels EM et al. (2007): Efficacy and safety of the weight-loss drug rimonabant: a meta-analysis of randomised trials. *Lancet*, Band 370, 1706–1713
- Curioni C, Andre C (2006): Rimonabant for overweight or obesity. *Cochrane Database, Systematic Review* CD006162
- Padwal R, Li SK, Lau DC (2004): Long-term pharmacotherapy for obesity and overweight. *Cochrane Database, Systematic Review* CD004094

bei der Therapie einzelner Patienten einzuschätzen, also der Situation am Krankenbett. Hier kann der Glauben an die Überlegenheit einer Therapie einen durchaus erwünschten Kontext-Faktor darstellen.

Bezogen auf das Beispiel Rimonabant (siehe Seite 17) war nach Meinung der Skeptiker die Datenlage nicht ausreichend, um eine positive Bewertung der Substanz treffen zu können. Es fehlten Studien mit klinischen Endpunkten und Aussagen über deren langfristigen Verlauf. Ebenso hätte bei der Interpretation der Daten die hohe Rate der Therapiebrecher deutlicher berücksichtigt werden müssen. Die in den Studien erlangten Ergebnisse führten so zu einer Überschätzung der positiven Effekte und aller Wahrscheinlichkeit nach zu einer Unterschätzung der Rate an unerwünschten Wirkungen. Im Ergebnis eine Situation, die als besonders gefährlich einzuschätzen ist, insbesondere dann, wenn therapeutische Alternativen bestehen und es sich nicht um eine unmittelbar lebensbedrohliche Erkrankung handelt. Inwieweit es sich bei Rimonabant um eine primär interessengetriebene Entscheidung gehandelt hat oder um eine Verletzung der Sorgfaltspflicht bei der Bewertung von Studien ist offen.

Bei jeder wissenschaftlichen Tätigkeit können bestimmte Interessenlagen identifiziert werden. Wichtig ist, sich über Folgendes bewusst zu sein: Je unmittelbarer die Entscheidungen finanzielle Auswirkungen haben, zum Beispiel in Form von Systementscheidungen, umso mehr ist auf eine möglicherweise interessengesteuerte Darlegung und Interpretation von Ergebnissen zu achten. Die Frage nach der Einhaltung erforderlicher Standards bei der Sorgfalt ist dabei auf jeder Ebene immer wieder zu stellen.

4 Evaluation: Der richtige Zeitpunkt

Wie oben bereits ausgeführt, sind zu einem frühen Zeitpunkt im Lebenszyklus einer Innovation Wunsch und Bedarf einer Entscheidung besonders groß. Dies gilt für Zulassungsentscheidungen in gleicher Weise wie für Entscheidungen über ihre Markteinführung beziehungsweise Erstattungsfähigkeit oder für die Frage nach den potenziellen finanziellen Auswirkungen für das Gesundheitssystem eines Landes. So ist auch zu erklären, warum zunehmend Publikationen zu den Fragen erscheinen, wann eine gesundheitsökonomische Analyse durchgeführt werden kann und wie bei „frühen“ Berechnungen mit dem Problem der bis dahin nicht vorhandenen Daten umgegangen werden kann. Das nicht banale Problem der Wahl des optimalen Zeitpunktes für eine Evaluation ist erst noch zu lösen. Buxton's Law lautet: "It is always too early to evaluate until, suddenly, it is too late" (nach *Hollingworth und Jarvik 2007*). Zu früh, da

valide Daten nicht vorliegen, zu spät, weil entweder bereits Schaden entstanden ist oder eine Evaluation, ethisch nicht mehr vertretbar erscheint. Das Problem ist bei sich schnell entwickelnden Technologien gravierend (*Hollingworth und Jarvik 2007*). Besonders groß ist das Problem des richtigen Zeitpunkts einer Nutzenbewertung häufig dann, wenn von kleinen Effekten bezogen auf die Überlegenheit einer Innovation gesprochen wird, denen aber möglicherweise gravierende unerwünschte Wirkungen gegenüberstehen, über die nur Erkenntnisse mit größerer Unsicherheit vorliegen.

Djulgovic hat als einen Ansatz für den Umgang mit dem Problem der Entscheidung bei unsicherer Datenlage den Begriff des „acceptable regret“ geprägt (*Hozo et al. 2008*). Er verdeutlicht dies an der Entscheidung darüber, ob Frauen im Alter zwischen 40 und 49 Jahren die Teilnahme an einem regelmäßigen Mammographie-Screening zu empfehlen sei (*Djulgovic und Lyman 2006*). Die Daten hinsichtlich des potenziellen Nutzens zeigen mit großer Verlässlichkeit einen kleinen Effekt, die Daten über den potenziellen Schaden sind wenig valide. Eine Aussage darüber, ob am Ende ein Netto-Nutzen resultiert, ist daher mit großer Unsicherheit behaftet. Es bleibt nur, im Einzelfall mit der Frau zu entscheiden, ob das Bedauern dann größer wäre, wenn ein Karzinom auftritt, welches vielleicht früher hätte diagnostiziert werden können oder dann, wenn möglicherweise später im Leben ein Karzinom in Folge des regelmäßigen Screenings auftritt (*Djulgovic und Lyman 2006*).

Das Mammographie-Beispiel verdeutlicht, dass die Entscheidung, egal ob individuell oder für das gesamte System, immer vom Kontext abhängt, zum Beispiel von Alternativen und vom Grundrisiko oder von Wertvorstellungen und Präferenzen. So stellt sich die Frage, in welchem Ausmaß mögliche Risiken bei Entscheidungen berücksichtigt werden sollten. Unabhängig von der Antwort auf diese Frage sollte einer Entscheidung immer die transparente Darlegung aller zur Verfügung beziehungsweise nicht zur Verfügung stehenden Daten zugrunde liegen. Wann immer möglich, sollten im Zuge von Sensitivitätsanalysen auch sogenannte Worst- und Best-case-Szenarien gerechnet werden.

Dies hätte im Fall von Rimonabant (siehe Seite 17) für die vorhandenen systematischen Übersichtsarbeiten geschehen müssen, die weitaus deutlicher auf die potenziellen Schäden hätten hinweisen müssen. Dass die systematischen Übersichtsarbeiten zu Rimonabant bei Adipositas trotz Marktrücknahme des Arzneimittels weiter unverändert sind, weist darauf hin, dass bei den Aktualisierungen der systematischen Übersichtsarbeiten der Cochrane-Library Mängel bestehen.

5 Frühe gesundheitsökonomische Analysen: Maximale Datentransparenz erforderlich

Hartz und John legten einen systematischen Übersichtsartikel zu dem möglichen Beitrag frühzeitiger gesundheitsökonomischer Analysen auf die weitere Produktentwicklung und gesundheitspolitische Entscheidungen vor (Hartz und John 2008). In die Arbeit sind 83 gesundheitsökonomische Analysen eingegangen, welche „early data“ verwendeten, weitere 41 Arbeiten, die die Rolle früher gesundheitsökonomischer Analysen aus Sicht der Hersteller diskutieren und 47 Arbeiten, die mögliche Konsequenzen aus Sicht gesundheitspolitischer Entscheidungsträger untersuchen. Leider ist die Auswertung narrativ, sodass hier nur über einzelne Aspekte berichtet werden kann. Einer eher positiven Grundeinstellung entsprechend gehen die Autoren davon aus, dass das Hauptdilemma gesundheitspolitischer Entscheider darin bestünde, das sie möglichst schnell gewährleisten wollten, dass allen Bürgern gleicher Zugang zu Innovationen gewährleistet werden soll, dieser Zugang aber unter Umständen durch limitierte Budgets eingeschränkt

sei. Vor diesem Hintergrund komme einer frühen gesundheitsökonomischen Bewertung der neuen Technologie besondere Bedeutung zu. Die Schwierigkeit einer begrenzten Datenlage berücksichtigend werden drei potenzielle Vorgehensweisen beschrieben:

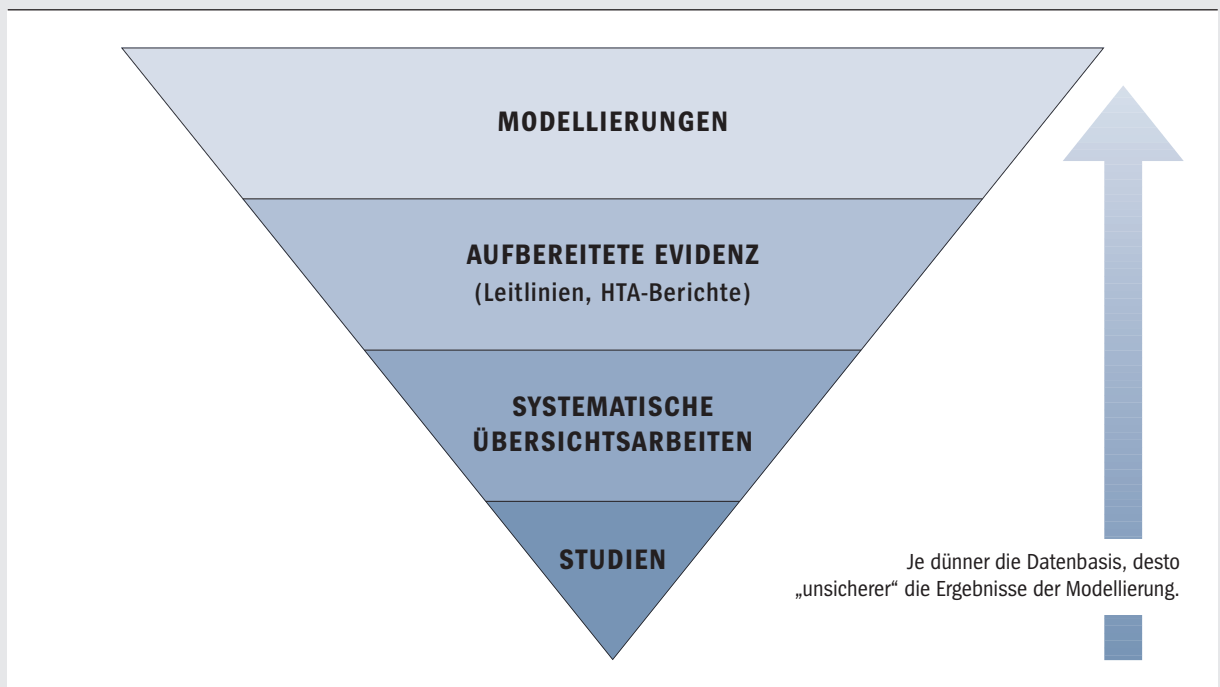
1. Zulassung (zum Markt) ohne Kenntnis der ökonomischen Auswirkungen;
2. keine Zulassung, solange keine valide Daten vorliegen;
3. bedingte Zulassung mit der Auflage, weitere Daten zu generieren.

Die Autoren konstatieren, dass es für den Umgang mit Unsicherheit, der sich automatisch aus dem Zeitpunkt einer frühen gesundheitsökonomischen Analyse ergibt, keine allgemein gültige Lösung gibt. Die Möglichkeit einer bedingten Zulassung beziehungsweise einer Zulassung mit Auflagen ermöglicht einen Weg zwischen klarem ja und nein (Kamrrow 2007).

Die Übersichtsarbeit von Hartz und John bietet letztlich keine Lösung in der Frage nach dem zu empfehlenden Umgang mit Unsicherheit aufgrund begrenzter Daten. Die

ABBILDUNG 1

Modifizierte Evidenzpyramide („auf Sand gebaut“)



Quelle: Leigemann 2009

Autoren plädieren jedoch dafür, alle Unwägbarkeiten transparent darzulegen.

Gerade hier scheint ein Problem gesundheitsökonomischer Analysen zu liegen. Eine Übersichtsarbeit, in der 13 Studien (verschiedene gesundheitsökonomische Modelle) zum Screening auf abdominelles Aortenaneurysma (AAA) analysiert wurden, kommt zu dem Ergebnis, dass den Publikationen keine ausreichenden Angaben zur Qualität der Modelle und den zugrunde gelegten Annahmen entnommen werden können (*Campbell et al. 2007*). Die Autoren vermuten, dass dies der Grund ist, dass die gesundheitsökonomischen Studien keine Beachtung bei der Entscheidung über die Einführung eines Screenings für AAA gefunden haben, obwohl die Ergebnisse überwiegend die Kosteneffektivität der Maßnahme nachgewiesen haben.

Bezogen auf das Fallbeispiel Rimonabant (siehe Seite 17) gelten die beschriebenen Mängel ebenfalls. Besonders erschwerend ist, dass die gesundheitsökonomische Modellierung nur in Form eines Abstracts publiziert ist, dem entscheidende Informationen nicht entnommen werden können. Eine Anfrage bei den Autoren (Datum 15.12.2008) ergab, dass keine Volltextpublikation existiert. Dies könnte Folge oder auch Ursache des „optimism bias“ sein. Die in Kongress-Abstracts dargestellten Ergebnisse von Studien überschätzen in relevantem Umfang die tatsächlichen Ergebnisse. Der „Fluch des Ruhms“ führt dazu, dass Limitationen eigener Arbeiten nicht ausreichend dargelegt werden (*Ioannidis 2007; Young et al. 2008*). Insbesondere kann nicht beurteilt werden, inwieweit auch mögliche unerwünschte Wirkungen berücksichtigt wurden und die Errechnung von Best- und Worst-case-Szenarien erfolgte. Die Tatsache, dass offensichtlich aufbauend auf dem gesundheitsökonomischen Modell Verträge mit Patienten geschlossen wurden, die die kontinuierliche Einnahme des Arzneimittels honorierten, ist besonders bitter. An dieser Stelle könnte man in Abwandlung eines bekannten Buchtitels (*Schmacke 2005*) fragen: „Wie viel Innovation verträgt der Mensch?“ Rimonabant ist ein Beispiel dafür, wie eine Innovation auf methodisch wackeligem Grund über unkritische Sekundärforschung auf einen Thron gehoben wird, der erst umstürzt, wenn die Beweislast des Schadens erdrückend wird (Abbildung 1).

6 Fazit

Die tatsächlichen Auseinandersetzungen um Innovationen und deren Bedeutung und Bewertung im Laufe der letzten zehn Jahre (seit Einführung von evidenzbasierter Medizin und Health Technology Assessment in Deutschland) haben sich nur unwesentlich geändert. Sie werden lediglich über

andere Instrumente auf methodisch höherem Niveau ausgetragen. Unterschiedliche Grundhaltungen und Werturteile kommen auch in Bewertungen nach streng formalisierten Verfahren mit rigider Methodik zum Ausdruck und beeinflussen deren Ergebnisse.

Das Dilemma, Entscheidungen treffen zu müssen, obwohl die Datenlage unsicher ist, ist durch die Anwendung der Prinzipien der evidenzbasierten Medizin deutlicher geworden, gelöst werden kann es so nicht. Je weiter die Verwendungen und Aufbereitungen von Evidenz sich von der zugrunde liegenden Datenbasis entfernen, umso höher sind die Anforderungen an die anzuwendende Transparenz und Sorgfalt. Das Prinzip des „acceptable regrets“ sollte immer berücksichtigt und Entscheidungen nur unter Kenntnis und expliziter Darlegung vorhandener Unsicherheiten getroffen werden.

Danksagung: Mein Dank gilt Herrn Professor Dr. Norbert Schmacke (Sprecher des Vorstands des HTA-Zentrums Bremen) für den Hinweis auf den „Fall Rimonabant“ und Herrn Martin Gerken (Wissenschaftlicher Mitarbeiter am HTA-Zentrum Bremen) für die Aufbereitung der Daten zu Rimonabant.

Literatur

Campbell H, Briggs A, Buxton M et al. (2007): The credibility of health economic models for health policy decision-making: the case of population screening for abdominal aortic aneurysm. *J Health Serv Res Policy*, Band 12, 11–7

Chalmers I, Matthews R (2006): What are the implications of optimism bias in clinical research? *Lancet*, Band 367, 449–450

Djulbegovic B, Kumar A, Soares HP et al. (2008): Treatment success in cancer: new cancer treatment successes identified in phase 3 randomized controlled trials conducted by the National Cancer Institute-sponsored cooperative oncology groups, 1955 to 2006. *Arch Intern Med*, Band 168, 632–642

Djulbegovic B, Lyman GH (2006): Screening mammography at 40–49 years: regret or no regret? *Lancet*, Band 368, 2035–2037

Donner-Banzhoff N, Schrappe M, Lelgemann M (2007): Studien zur Versorgungsforschung. Eine Hilfe zur kritischen Rezeption. *Z Ärztl Fortbild Qualitatssich*, Jg. 101, 463–471

Hagen A, Gerhardus A (2008): 9. Symposium Health Technology Assessment, Köln 2008 (Abstract)

Hartz S, John J (2008): Public health policy decisions on medical innovations: What role can economic evaluation play. *Health Policy* 2008, elektronische Publikation: doi:10.1016/j.healthpol.2008.05.011

Hollingworth W, Jarvik JG (2007): Technology assessment in radiology: putting the evidence in evidence-based radiology. *Radiology*, Band 244, 31–38

Hozo I, Schell MJ, Djulbegovic B (2008): Decision-making when data and inferences are not conclusive: risk-benefit and acceptable regret approach. *Semin Hematol*, Band 45, 150–159

Ioannidis JP (2007): Limitations are not properly acknowledged in the scientific literature. *J Clin Epidemiol*, Band 60, 324–329

Kamerow D (2007): Paying for promising but unproven technologies. *BMJ*, Band 335, 965

Kumar A, Soares H, Wells R et al. (2005): Are experimental treatments for cancer in children superior to established treatments? Observational study of randomised controlled trials by the Children's Oncology Group. *BMJ*, Band 331, 1295

Schmacke N (2005): Wie viel Medizin verträgt der Mensch? Bonn: KomPart

Strech D (2007): Vier Ebenen von Werturteilen in der medizinischen Nutzevaluation. Eine Systematik zur impliziten Normativität in der Evidenz-basierten Medizin. *Z Ärztl Fortbild Qualitatssich*, Jg. 101, 473–480

Young NS, Ioannidis JP, Al-Ubaydli O (2008): Why current publication practices may distort science. *PLoS Med*, Band 5, e201

DIE AUTORIN

**Monika Lelgemann, MSc,**

Jahrgang 1959, Fachärztin für Anästhesie und klinische Epidemiologin, studierte an den Universitäten Marburg und Göttingen und Rotterdam. Nach langjähriger Tätigkeit am Institut für Klinische Pharmakologie in Bremen arbeitete sie von 2001 bis 2004 als wissenschaftliche Mitarbeiterin am Deutschen Cochrane Zentrum in Freiburg. In den Jahren 2004 bis 2006 leitete sie

den Bereich Evidenzbasierte Medizin und Leitlinien am Ärztlichen Zentrum für Qualität in der Medizin in Berlin. Seit 2006 hat sie die wissenschaftliche Geschäftsführung des interdisziplinären HTA-Zentrums an der Universität Bremen inne. Zu ihren Arbeitsschwerpunkten gehört die Anwendung der Prinzipien der Evidenzbasierten Medizin im Zusammenhang mit der Erstellung von Leitlinien und HTA-Berichten.