www.monitor-versorgungsforschung.de KONGRESS-SPECIAL (1/2013) 6. JAHRGANG

VERSORGUNGS monitor FORSCHUNG

Fachzeitschrift zu Realität, Qualität und Innovation der Gesundheitsversorgung

Kongress-Special 1/2013

Der Endpunkt

Patienten Relevante Endpunkte

Wer definiert, was für den Patienten relevant ist?



„Trugschlüsse und Konsequenzen“ (Müller)

„Bewertung aus der Perspektive der Patienten“ (Mühlbacher)

„Conjoint Analyse versus Analytic Hierarchy Process? (Dintsios)

„Surrogate und Endpunkte beim Prostatakarzinom“ (Stark)

KONGRESS-SPECIAL „Der Endpunkt“

ISSN 1866-0533

VERSORGUNGS monitor FORSCHUNG

01/13 6. Jahrgang

KongressSpecial 1/2013

Wissenschaftlicher Beirat

Prof. Dr. B. Borisch MPH FRCPath / Prof. Dr. G. Glaeske / Dr. C. Hermann / F. Knieps / R. Lederer / Prof. Dr. W.-D. Ludwig / Prof. Dr. E. A. M. Neugebauer / Prof. Dr. h.c. H. Rebscher / Dr. J. Roski MPH / Prof. Dr. M. Schrappe / Dr. T. Trümper

Praxisbeirat

arvato services healthcare/ Deutsche BKK/Fresenius Kabi/Kassenärztliche Vereinigung Bayerns/Novartis Pharma GmbH/MedicalContact AG/Pfizer Deutschland GmbH/ Vivantes - Netzwerk für Gesundheit GmbH



Nutzen für den Patienten

Dieses Kongress-Special 1/2013 enthält die ersten Originalbeiträge zum 3. Fachkongress von Monitor Versorgungsforschung, den wir am 29. Oktober 2012 in Berlin zusammen mit der Deutschen Gesellschaft für bürgerorientiertes Versorgungsmanagement (DGbV) durchgeführt haben. Unter dem Thema Patientenrelevante Endpunkte ging es um das General-Ziele der Gesundheitsversorgung: Was kommt wirklich beim Patienten an? Wie kommt man methodisch von Surrogatparametern zum tatsächlichen Nutzen für den Patienten?

Auf den folgenden beiden Seiten haben Sie einen Überblick über die Referenten und Vorträge. In diesem Special finden Sie die ersten Originalbeiträge dazu:

Müller und **Schaefer** analysieren, wie man von Surrogatparametern und klinischen Endpunkten zu patientenrelevanten Endpunkten kommt, und weisen insbesondere auf Trugschlüsse hin, die dabei vermieden werden müssen.

Mühlbacher, Bethke und **Juhnke** beschäftigen sich mit der Vermittlung von Patientenpräferenzen mit der Methode der Discrete-Choice-Experimente.

Dintsios vergleicht die zwei Verfahren zur Erhebung von Patientenpräferenzen Conjoint Analyse (CA) und Analytic Hierarchy Process (AHP) nach ihren Ergebnissen, Vor- und Nachteilen.

Stark untersucht die Indikation Prostatakarzinom und fordert für künftige Neuentwicklungen von Medikamenten die Entwicklung von Surrogatparametern, die prädiktiv für das Gesamtüberleben und das Ansprechen auf Therapien sind.

Ich wünsche Ihnen eine interessante und spannende Lektüre!

Mit herzlichen Grüßen

Ihr

Professor Dr. Reinhold Roski

Hardy Müller / Corinna Schaefer 5

Surrogatparameter und patientenrelevante Endpunkte: Trugschlüsse und Konsequenzen

Surrogate sind Ersatz-Messparameter, die eine vom Patienten nicht unmittelbar wahrnehmbare Vorstufe eines klinischen Endpunktes darstellen. In der klinischen Forschung werden sie häufig eingesetzt, weil sie weniger aufwändig zu erheben sind als klinische Endpunkte. Damit ein Surrogatparameter verlässliche Aussagen erlaubt, muss er validiert werden. Dazu muss u.a. in RCTs nachgewiesen werden, dass ein therapeutischer Effekt auf den Surrogatparameter einen kausalen Effekt auf den klinischen Zielparameter hat. Die meisten heute genutzten Surrogate sind jedoch nicht validiert. Ihr Einsatz hat immer wieder zu Trugschlüssen geführt, d.h. beobachtete Effekte auf den Surrogatendpunkt hatten keine oder gegenteilige Auswirkungen auf den Zielendpunkt. Behandlungs- und Systementscheidungen im Gesundheitswesen sollen daher nicht auf der Basis von Surrogaten getroffen werden, in der Praxis geschieht das jedoch häufig.

Prof. Dr. Axel C. Mühlbacher / Susanne Bethge, M.Sc. / Christin Juhnke, M.A. 11

„Patientenrelevante Endpunkte" - Bewertung aus der Perspektive der Patienten?

Der dokumentierte Patientennutzen sollte ausschlaggebend für die Prioritätensetzung sein. Die zugrundeliegenden Entscheidungsprozesse müssen dafür Sorge tragen, dass die Auswahl aus einer Vielzahl potentieller Gesundheitsgüter und -dienstleistungen mit den Bedürfnissen, Erwartungen und Prioritäten der Versicherten und Patienten im Einklang stehen. Die gemeinsame Basis des klinischen und des ökonomischen Entscheidungskalküls ist der Nutzen für die Patienten. Patientenpräferenzen sollten deshalb auf allen Ebenen in die Entscheidungsprozesse einbezogen werden. Zur Präferenzmessung im Gesundheitswesen werden zunehmend Discrete-Choice-Experimente eingesetzt.

Dr. rer. biol. hum. Ch.-M. Dintsios, MPH 16

Patientengewichtete Endpunkte: Conjoint Analyse versus Analytic Hierarchy Process

Beide Verfahren, Conjoint Analyse (CA) und Analytic Hierarchy Process (AHP), werden zur Erhebung von Patientenpräferenzen (hier Priorisierung und Gewichtung von Endpunkten) eingesetzt. Über ihre theoretische Verortung hinaus (ökonomische Nutzentheorie versus deskriptive Theorie) unterscheiden sie sich hinsichtlich des zugrundegelegten Modells (additives Teilwertmodell versus gewichtetes additives Nutzenmodell), methodischen Ansatzes (dekompositionell versus kompositionell), Bewertungsprozesses (holistisch versus Paarvergleiche), methodischer Schwächen (kognitive Überforderung versus unrealen Entscheidungssituationen) und des Ressourcenaufwands (umfangreich versus gering) nicht unerheblich.

Dr. med. Thomas Stark 21

Surrogate und Endpunkte beim Prostatakarzinom

Mit der Verfügbarkeit neuer Therapiemöglichkeiten hat sich die Prognose für Patienten mit metastasiertem kastrationsresistenten Prostatakarzinoms (mCRPC) deutlich verbessert. Aufgrund der Entwicklungszyklen ergibt sich für zukünftige Neuentwicklungen von Medikamenten die Notwendigkeit, Surrogatendpunkte zu definieren, die prädiktiv für das Gesamtüberleben und ein Therapienansprechen sind und gleichzeitig von den Gesundheitsbehörden anerkannt werden.

Impressum **Monitor Versorgungsforschung** - Fachzeitschrift zu Realität, Qualität und Innovation der Gesundheitsversorgung

Monitor Versorgungsforschung
Fachzeitschrift zu Realität, Qualität und Innovation der Gesundheitsversorgung
6. Jahrgang/ISSN: 1866-0533

Herausgeber
Prof. Dr. Reinhold Roski, Berlin
roski@m-vf.de

Chefredaktion
Peter Stegmaier (verantw. Redakt.)
Kölnstr. 119, 53111 Bonn
Tel +49 228 7638280-0
Fax +49 228 7638280-1
stegmaier@m-vf.de

Redaktion
Wolfgang Dame
dame@m-vf.de
Olga Gnedina
gnedina@m-vf.de
Jutta Mutschler
mutschler@m-vf.de

Verlag
eRelation AG - Content in Health
Vorstand: Peter Stegmaier
Kölnstr. 119, 53111 Bonn
www.erelation.org
mail@erelation.org

Verlagsleitung
Peter Stegmaier

Anzeigenleitung/Vertrieb/Abo
Anke Heiser (verantwortlich für den Anzeigenteil)
heiser@m-vf.de

Marketing:
Kölnstr. 119, 53111 Bonn
Tel +49 228 7638280-0
Fax +49 228 7638280-1

Abonnement
„Monitor Versorgungsforschung" erscheint sechsmal jährlich. Der Preis für ein Jahresabonnement beträgt 90 Euro. Jahresvorzugspreis für Studenten gegen Vorlage einer Immatrikulationsbescheinigung 60 Euro. Die genannten Preise verstehen sich zzgl. Versandkosten: Inland 9,21 Euro; Ausland 36 Euro. Preisänderungen vorbehalten. Die Abonnementdauer beträgt ein Jahr. Das Abonnement verlängert sich automatisch um ein weiteres Jahr, wenn es nicht spätestens sechs Wochen vor Ablauf des Bezugsjahres schriftlich gekündigt wird.

Layout
eRelation AG, Bonn

Druck
Kössinger AG
Fruehaufstraße 21
84069 Schierling
info@koessinger.de
Tel +49-(0)9451-499124
Fax +49-(0)9451-499101
Printed in Germany

Urheber- und Verlagsrecht
Die Zeitschrift und alle in ihr enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Mit Annahme des Manuskripts gehen das Recht zur Veröffentlichung sowie die Rechte zur Übersetzung, zur Vergabe von Nachdruckrechten, zur elektronischen Speicherung in Datenbanken, zur Herstellung von Sonderdrucken, Fotokopien und Mikrokopien an den Verlag über. Jede Verwertung außerhalb der durch das Urheberrechtsgesetz festgelegten Grenzen ist ohne Zustimmung des Verlags unzulässig. In der unaufgeforderten Zusendung von Beiträgen und Informationen an den Verlag liegt das jederzeit widerrufliche Einverständnis, die zugesandten Beiträge bzw. Informationen in Datenbanken einzustellen, die vom Verlag oder Dritten geführt werden.

Auflagenmeldung
Mitglied der Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V. (IVW), Berlin. Verbreitete Auflage: 6984 (IVW 4. Quartal 2012)

Die Vortragenden des Kongresses „Patientenrelevante Endpunkte“			
Thema	**Wer**		**Abstract**
Was ist für den Patienten relevant? - Die eigene Sicht und Position der Patienten	Dr. Michael Köhler, Patientenliga Atemwegserkrankungen e.V		Mortalität ist unabänderbar, Morbidität oft nicht, Laborrelevanz ist nicht gefühlt relevant, gemessene Zeit ist nicht gefühlte Zeit. Versuch lebensrelevante Kriterien am Beispiel der COPD zu erfassen. Weg von alleinigen „objektiven“ Messparametern zur individualisierten Symptomerfassung und Krankheitsbewertung.
Widersprüche zwischen Surrogaten und patientenrelevanten Endpunkten: Trugschlüsse und Konsequenzen **siehe Seite 5**	Hardy Müller M.A., WINEG		Surrogate sind Ersatz-Messparameter, die eine vom Patienten nicht unmittelbar wahrnehmbare Vorstufe eines klinischen Endpunktes darstellen. In der klinischen Forschung werden sie häufig eingesetzt, weil sie weniger aufwändig zu erheben sind als klinische Endpunkte. Damit ein Surrogatparameter verlässliche Aussagen erlaubt, muss er validiert werden.
Was wird als patientenrelevant angesehen und was nicht?	Karen Pottkämper M.A., AQUA Institut		
Kosten-Nutzen-Bewertung zwischen Anspruch und Datenbasis: Surrogate und patientenrelevante Endpunkte	PD Dr. Andreas Gerber-Grote, IQWIG		Für Entscheidungen im System der GKV sollten nach SGB V patientenrelevante Endpunkte und keine sog. Surrogate herangezogen werden. Bei der Auswahl der Erhebungsverfahren müssen insbesondere Kriterien wie Validität, Repräsentativität und theoretische Fundierung berücksichtigt werden. Anhand von zwei Pilotprojekten zu Endpunkten bei Depression und Hepatitis C, die am IQWiG 2010 durchgeführt wurden, werden die Fragen exemplarisch diskutiert.
Patientendefinierte Endpunkte Pilotstudie Conjoint Analyse (CA) **siehe Seite 11**	Prof. Dr. Axel Mühlbacher, Hochschule Neubrandenburg, IGM		Das Methodenkonzept für die Kosten-Nutzen Bewertung sieht die Verwendung einer Effizienzgrenze innerhalb eines Indikationsgebietes vor. Der Nutzen wird vom IQWiG anhand patientenrelevanter Endpunkte bewertet. Bei der Entscheidungsfindung auf Basis mehrerer Effizienzgrenzen bleibt offen, wie die Ergebnisse durch den Entscheidungsträger interpretiert werden. Es sollte dargestellt werden, wie die Conjoint / Discrete-Choice Analyse zur Identifikation patientenrelevanter Endpunkte genutzt werden kann. Weiterhin sollte geprüft werden, ob diese Methode für die Gewichtung und Priorisierung über multiple Endpunkte anwendbar ist. Darauf aufbauend wurde geprüft, inwiefern diese Methode zur Priorisierung von Gesundheitstechnologien auf Basis multipler Endpunkte eingesetzt werden kann. Ein indikationsspezifischer Vergleich von Therapiealternativen ist denkbar. Zielsetzung der Machbarkeitsstudie waren die Identifikation und Gewichtung von Attributen am Beispiel der antiviralen Hepatitis C-Therapie. Im Ergebnis zeigt sich, dass die Discrete-Choice Methode eine theoretisch fundierte Methode zur Identifikation, Gewichtung und Priorisierung von patientenrelevanten Endpunkten bietet. Weiterhin können mit Hilfe von DCEs auch Präferenzunterschiede zwischen Experten und Patienten aufgezeigt und verglichen werden. Ferner kann die Wertfunktion approximiert und ein übergeordneter Nutzenwert gebildet werden. Die grafische Ableitung einer übergeordneten Effizienzgrenze ist somit darstellbar.
Patientendefinierte Endpunkte Pilotstudie AHP	Prof. Maarten J. IJzerman, PhD, University of Twente		In recent years Health Technology Assessment (HTA) processes have expanded beyond those countries dominated with a single payer, and with this expansion there has been renewed attention placed on the fundamental evaluation methods that are used. While there has been considerable attention placed on comparative effectiveness research (CER) in the US, another fundamental issue has emerged: how can HTA better involve patients. Here patient centered strategies and methods can be categorized into two types: those that involve patients directly in decision making and those that aim at the study of patients needs and preferences in terms of benefit-risk trade-offs. In determining these decision trade-offs, one may identify two subsequent steps, (1) the quantification of benefits and risks using the clinical evidence and (2) the judgment of these benefits and risks in terms of their relative preference. Patient preferences may be relevant in HTA by combining benefits and risks into one weighted aggregate measure (for drug approval), but also for incorporating patients' views in the reimbursement decisions. his presentation will discuss the process of benefit-risk assessment and discusses (dis-) advantages of preference elicitation methods, like multi-criteria decision analysis and analytic hierarchy process
Patientendefinierte Endpunkte: CA vs. Analytic Hierarchy Process (AHP) **siehe Seite 16**	Dr. Charlabos-Markos Dintsios, vfa		Beide Verfahren werden zur Erhebung von Patientenpräferenzen (hier Priorisierung und Gewichtung von Endpunkten) eingesetzt. Über ihre theoretische Verortung hinaus (ökonomische Nutzentheorie vs. deskriptive Theorie) unterscheiden sie sich hinsichtlich des zugrundegelegten Modells (additives Teilwertmodell vs. gewichtetes additives Nutzenmodell), methodischen Ansatzes (de-kompositionell vs kompositionell), Bewertungsprozesses (holistisch vs Paarvergleiche), methodischer Schwächen (kognitive Überforderung vs unrealen Entscheidungssituationen) und dem Ressourcenaufwand (umfangreich vs gering) nicht unerheblich. CA kann bezogen auf a priori Annahmen zu widersprüchlichen Ergebnissen führen, AHP leidet unter dem so genannten Rank Reversal. Sie generieren allerdings in vergleichenden Studien durchaus ähnliche Ergebnisse und besitzen je nach Fragestellung ihre Vor- und Nachteile.

Die Vortragenden des Kongresses „Patientenrelevante Endpunkte“			
Thema	**Wer**		**Abstract**
Was ist für den Patienten relevant? Beispiel Indikation Diabetes	Dr. med. Jens Kröger, DDG/diabetes.de		Die Stärkung des Menschen mit Diabetes hinsichtlich seiner Wünsche/Erwartungen im Rahmen der Behandlung ist ein wichtiges Handlungsfeld hinsichtlich der Verbesserung der Versorgung. Dabei divergieren nicht selten die Wünsche der Patienten mit wissenschaftlichen Forschungsfeldern. Im Rahmen eines erforderlichen nationalen Diabetesplanes sollten daher Maßnahmen entwickelt werden, um Patienten eine gezielte Unterstützung zu geben, auf der Basis selbstbestimmter Entscheidungen besser mit dem Diabetes und seiner Folgen umzugehen sowie soziale wie ökonomische Folgen der Erkrankung zu vermeiden.
Endpunkte in der Versorgung aus Sicht der Kassen	Prof. Dr. h.c. Herbert Rebscher, DAK		In § 35b SGB V heißt es zur Kosten-Nutzen-Bewertung von Arzneimitteln: „Beim Patienten-Nutzen sollen insbesondere die Verbesserung des Gesundheitszustandes, eine Verkürzung der Krankheitsdauer, eine Verlängerung der Lebensdauer, eine Verringerung der Nebenwirkungen sowie eine Verbesserung der Lebensqualität … angemessen berücksichtigt werden.“ Grundsätzlich ist aus Kassensicht an nichtmedikamentöse Therapien der gleiche Maßstab anzulegen.
Surrogate und Endpunkte bei Diabetes	Dr. Thomas M. Zimmermann, Lilly		An zahlreichen Stellen des SGB V wird auf die evidenzbasierte Medizin verwiesen, nach der die Patientenpräferenzen eine wesentliche Komponente in der Gesundheitsversorgung bilden. Dabei stellen sich die zentralen Fragen: Was wissen wir über die Patientenpräferenzen? Wer entscheidet eigentlich darüber, welche Endpunkte patientenrelevant sind? Bereits mit der Kosten-Nutzen-Methodik des IQWiG, spätestens jedoch mit der Verabschiedung des AMNOG hat die Diskussion um patientenrelevante Endpunkte an Fahrt aufgenommen. Weitgehender Konsens dürfte darüber bestehen, dass patientenrelevante Endpunkte und Präferenzen indikationsspezifisch ermittelt und betrachtet werden sollten. Für den Bereich Diabetes bedeutet dies minimal eine differenzierte Betrachtung der Präferenzen von Patienten mit DM Typ 1 und DM Typ 2. Neben einer grundlegenden Auseinandersetzung mit dem Thema werden Erkenntnisse aus zwei Präferenzstudien zu den beiden vorgenannten Diabetes-Typen vorgestellt. Surrogatendpunkte wie das radiographische PFS (rPFS), PSA-Response und Patient Reported Outcomes wie Lebensqualität (QoL), Schmerzlinderung etc. sind gegenwärtig von der Europäischen Zulassungsbehörde ebenfalls akzeptierte Endpunkte. Dennoch setzt auch eine Verwendung dieser Endpunkte die Annahme voraus, dass es sich bei mCRPC-Patienten um ein homogenes Patientenkollektiv handelt, welches mit einem globalen Endpunkt zufriedenstellend erfasst werden kann. Zugleich wird aber zunehmend eine personalisierte Medizin gefordert. Diese setzt jedoch voraus, dass unterschiedliche Patientenkollektive existieren, die sich naturgemäß in ihrer Prognose unterscheiden, sich aber damit auch der Beurteilung anhand eines globalen Endpunktes entziehen. Hier sind prädiktive Biomarker gefordert, die bislang aber für das mCRPC (noch) nicht ausreichend definiert bzw. validiert sind. Gleason-Score, PSA-Verdopplungszeit, Zeit bis zum Auftreten des Rezidivs bzw. zur Metastasierung, zirkulierende Tumorzellen (CTC) und Androgenrezeptormutationen werden als prognostische und ggf. auch prädiktive Marker zur individuellen Steuerung der Therapie diskutiert und könnten als Surrogatmarker für die Biologie des Prostatakarzinoms ein Einstieg in diese Diskussion sein. Die Validierung und Verwendung dieser prognostischen bzw. prädiktiven Marker, die die Tumorbiologie direkt oder indirekt widerspiegeln, könnte uns in Zukunft eine personalisierte, individualisierte Therapie und eine risikostratifizierte Behandlung von mCRPC-Patienten ermöglichen. Mit der Umsetzung einer personalisierten Medizin wäre aber auch verbunden, die in Studien verwendeten Endpunkte immer auch in Abhängigkeit einer risikoadaptierten Patientenstratifizierung zu betrachten. Hier sind gemeinsame Anstrengungen von Industrie, wissenschaftlicher Community und Behörden notwendig, um im Sinne der Patienten zeitnah validierte Surrogate zu etablieren und in der Bewertung klinischer Studien ein flexibles Bewertungssystem einzuführen, das der jeweiligen Biologie des Tumors und dem entsprechenden Surrogat gerecht wird.
Endpunkte aus Sicht der Ärzteschaft	Dr. Dominik Graf von Stillfried, ZI		
Surrogate und Endpunkte beim Prostatakarzinom **siehe Seite 21**	Dr. Thomas Stark, Janssen-Cilag		Allgemein gilt das Overall Survival (OS) als härtester Endpunkt in Studien zur Behandlung des Prostatakarzinoms. Therapien sind danach zu bewerten, inwieweit sie das Gesamtüberleben der Prostatakarzinompatienten verlängern. Lag in der Vergangenheit das mediane Überleben beim metastasierten kastrationsresistenten Prostatakarzinom (mCRPC) noch bei 1-3 Jahren, so zeigen neue Therapieansätze eine deutliche Verbesserung des OS. Damit wäre theoretisch eine mediane Nachbeobachtungszeit von 4-5 Jahren erforderlich, um überhaupt eine Verbesserung dieses Endpunkts zeigen zu können. Aufgrund der Entwicklungszyklen neuer Medikamente ist dies aber zum Zeitpunkt der Zulassung nicht möglich, so dass sich die Notwendigkeit ergibt, Surrogatendpunkte zu definieren, die hinreichend prädiktiv für das OS sind und gleichzeitig von den Gesundheitsbehörden anerkannt werden.

Hardy Müller
Corinna Schaefer

Surrogatparameter und patientenrelevante Endpunkte: Trugschlüsse und Konsequenzen

Zusammenfassung

Surrogate sind Ersatz-Messparameter, die eine vom Patienten nicht unmittelbar wahrnehmbare Vorstufe eines klinischen Endpunktes darstellen. In der klinischen Forschung werden sie häufig eingesetzt, weil sie weniger aufwändig zu erheben sind als klinische Endpunkte. Damit ein Surrogatparameter verlässliche Aussagen erlaubt, muss er validiert werden. Dazu muss u.a. in RCTs nachgewiesen werden, dass ein therapeutischer Effekt auf den Surrogatparameter einen kausalen Effekt auf den klinischen Zielparameter hat. Die meisten heute genutzten Surrogate sind jedoch nicht validiert. Ihr Einsatz hat immer wieder zu Trugschlüssen geführt, d.h. beobachtete Effekte auf den Surrogatendpunkt hatten keine oder gegenteilige Auswirkungen auf den Zielendpunkt. Behandlungs- und Systementscheidungen im Gesundheitswesen sollen daher nicht auf der Basis von Surrogaten getroffen werden, in der Praxis geschieht das jedoch häufig. Als Konsequenz aus dieser unhaltbaren, die Patientensicherheit gefährdenden Situation sind folgende Forderungen abzuleiten: Verbindliche Vorgaben hinsichtlich klinischer Zielendpunkten für die Zulassung von Medikamenten und Standards für klinische Studien, um Daten zu patientenrelevanten Endpunkten zu generieren; konsequente Patientenbeteiligung auf allen Entscheidungsebenen des Gesundheitssystems, um Entscheidungen auf der Basis patientenrelevanter Endpunkte sicherzustellen. Vor allem ist es dringend notwendig die anstehenden Entwicklungen und Entscheidungen im Gesundheitssystem daraufhin zu prüfen, ob die ihnen zugrundeliegenden Annahmen auf validen Endpunkten beruhen.

Schlüsselwörter

Surrogatendpunkte, Surrogatparameter, Patientenrelevante Endpunkte, Nutzen, Partizipative Entscheidungsfindung, Personalisierte Medizin, Patientenrechte, Prävention

Die medizinische Versorgung wird in Deutschland zum größten Teil mit Versichertenbeiträgen finanziert. Daraus resultiert die normativ-ethisch begründete Verpflichtung der gesetzlichen Krankenversicherung, ein Optimum an Gesundheitsleistungen mit maximalem Nutzen für die Versicherten zu realisieren. Verschwendung gilt es zu reduzieren, Ressourcen effizient einzusetzen, alle Ergebnisse müssen insbesondere für die Patientinnen und Patienten bedeutsam sein. Das Erreichen dieser Ziele und ein primär wissenschaftlich fundiertes Vorgehen verantworten die Krankenversicherungen zusammen mit den Leistungserbringern und den Patienten (SGB V §§ 2, 12, 35b, 70). Die Frage, was nützlich sei, ist z.B. im SGB V § 35b beschrieben als „Verbesserung des Gesundheitszustandes, Verkürzung der Krankheitsdauer, eine Verlängerung des Lebensdauer, Verringerung der Nebenwirkungen sowie Verbesserung der Lebensqualität“. Diese Maxime findet sich in inhaltlich identischer Form in der Verfahrensordnung des G-BA, in der Arzneimittel-Nutzenbewertungs-Verordnung oder im Methodenpapier zur Nutzenbewertung des IQWiG. Besondere Bedeutung hat die Feststellung, dass sich die Nutzenbewertung an „patientenrelevanten Zielgrößen“ zu orientieren hat. „Als ‚patientenrelevant‘ soll in diesem Zusammenhang verstanden werden, wie ein Patient fühlt, seine Funktionen und Aktivitäten wahrnehmen kann oder ob er überlebt.“ (IQWiG 2011, S. 30f)

>> Mit diesen Definitionen sind die Ziele auch neuer Behandlungsverfahren und die Parameter einer Nutzenbewertung umrissen. Die Ziel-Kriterien sind oftmals nicht einfach oder schnell zu erheben, sondern erfordern wissenschaftlich verlässliche Nachweise. Die Erhebung kann teuer sein - es werden dafür Studien in aufwändigen Designs benötigt. Zudem erfordert die Bemessung patienentenrelevanter Endpunkte Zeit, die bei drängendem Handlungsbedarf bei Erkrankungen mit hoher Letalität ohne hinreichende Behandlungsoptionen (scheinbar) nicht zur Verfügung steht.

Zur Lösung dieser Probleme und raschen Bewertung der Ergebnisse medizinischer Behandlung werden Surrogatparameter genutzt. Diese Surrogate dienen als Ersatzkenngrößen für die eigentlichen klinischen Ziele. Sie sollen früher und einfacher zu erheben sein und die eigentlichen Zielgrößen angemessen repräsentieren.

Allen Surrogaten gemein ist (Mangiapane, Garrido 2009: 14), dass sie i. d. R. vom Patienten nicht unmittelbar wahrnehmbare Vorstufen des eigentlich klinischen Endpunktes darstellen. Häufig werden hierfür Biomarker genutzt, wie etwa das glykosylierte Hämoglobin (HbA1c) oder das LDL-Cholesterin. Die Korrelation zwischen Surrogat und klinischem Endpunkt (individual-level association, Buyse 2009) wie auch der kausale Zusammenhang zwischen Behandlungeffekten auf Surrogate und Effekten auf die klinischen Endpunkte (trial level association, ebd.) wird biologisch oder pathophysiologisch erklärt. Surrogate gelten dann als valide, wenn Assoziation und Kausalität verlässlich nachgewiesen wurden. Diese Validierung der Surrogate in jedem ihrer Einsatzgebiete ist eine eigenständige, unverzichtbare Aufgabe, die „... ausführliche Forschung, einschließlich der Durchführung von randomisiert kontrollierten Studien (RCT) mit klinisch relevanten Endpunkten [erfordert]“ (Mangiapane, Garrido 2009: 1).

Die Erfahrungen mit Surrogat-Endpunkten

Als Beispiel für einen sinnvollen Einsatz eines Surrogat-Parameters wird die Beurteilung der Wirksamkeit der antiviralen Therapie (AIDS/HIV) anhand der Anzahl von CD-4 Zellen und dem Nachweis von Virus RNA genannt (Glaeske 2010: 162). Diese Parameter korrelieren mit dem Fortschreiten der Erkrankung und boten die Grundlage für die Zulassung des Medikamentes durch die amerikanische Arzneimittelbehörde. Dadurch wurde eine schnelle Zulassung möglich, die vielen Menschen mit AIDS das Leben gerettet hat und damit effektiver war, als auf das Auftreten von opportunistischen Infektionen bzw. Todesfällen zu warten.

Doch Erfahrungen mit Surrogat-Endpunkten mahnen zu Vorsicht und größter Skepsis. Zu oft haben sich in der Vergangenheit an Surrogaten vermeintlich erfolgreiche Interventionen im Lauf der Zeit nach weitergehender Untersuchung nicht nur als unnütz, sondern leider oft auch als schädlich herausgestellt.

Daher ist die Anwendung von Surrogaten in den letzten Jahren zunehmend in die Kritik geraten. Studien, die ausschließlich Surrogate berichten, sind wissenschaftlich kaum zu rechtfertigen. Dies belegen eindringlich die so genannten Surrogattrugschlüsse (Mühlhauser, Berger 1996; Mühlhauser, Meyer 2006). Davon spricht man, wenn medizinische Interventionen zwar erwünschte Effekte auf das Surrogat haben, jedoch keine oder sogar gegenteilige auf die für Patienten relevanten Endpunkte. Beispiele hierfür finden sich in Tabelle 1.

Der HTA-Bericht „Surrogatendpunkte als Parameter der Nutzenbewertung“ (Mangiapane, Garrido 2009) zeigt, dass Surrogate in Health Technology Assessment (HTA) Berichten zur Bewertung der Relevanz von medizinischen Maßnahmen weltweit so gut wie nicht benutzt werden und deren Anwendung auch nicht empfohlen wird.

Methodenvergleich CA versus AHP				
Präventionsziel (Endpunkt)	Surrogat-Parameter	Intervention	Surrogat-Ergebnis	relevante, definitive Endpunkt-Ergebnisse
Menopausale Hormontherapie				
Prävention von Herzinfarkt	Cholesterin-/Blutzucker-Spiegel	Östrogene, Gestagene	Senkung von Cholesterin- und Blutzuckerspiegel	Zunahme von Herzinfarkt, Schlaganfall, Lungenembolie
Osteoporose				
Verminderung osteoporotischer Frakturen	Knochendichte (Densitometrie)	75 mg NaF plus 1500 mg Calcium pro Tag vs Placebo und 1500 mg Calcium pro Tag	Zunahme der Knochendichte	Zunahme der Extravertebralfrakturen um (72 vs 24, p < 0,01) in der NaF-Gruppe
Anämie				
Verbesserung von Leistungsfähigkeit bei Krebserkrankungen	Hämoglobinwerte	Erythropoietin	Zunahme der Hämoglobinwerte	Zunahme der Mortalität
Vitamine				
Verhinderung von Lungenkrebs	Vitamin-Serumspiegel	Vitamin E 50 mg pro Tag oder Beta-Caroten 20 mg pro Tag	Anstieg der Serum-Vitaminspiegel	Unter Beta-Caroten Anstieg von Lungenkrebs um 18%, Gesamtmortalität um 8%
Prävention von Krebs	Vitamin-Serumspiegel	Hochdosiert Vitamin E	Anstieg des Serum-Vitaminspiegels	Zunahme der Gesamtmortalität
Kardiovaskuläre Risikofaktoren				
Prävention von koronarer Herzkrankheit und vorzeitigem Tod bei Diabetes	Blutzucker	Rosiglitazone	Senkung des Blutzuckerspiegels	Zunahme von Herzinfarkten und Herzinsuffizienz

Tab. 1: Surrogattrugschlüsse (Ausschnitt aus Mühlhauser, Müller 2009, siehe dort auch Referenzen).

Diese schlechten Erfahrungen mit der Bemessung medizinischer Effekte auf der Basis von Surrogaten und die eingangs genannte Verpflichtung zu einer primär wissenschaftlich begründeten, schadensminimierten Behandlung erfordern zukünftig den weitest gehenden Verzicht auf Surrogate und die ausschließliche Bewertung medizinischer Verfahren auf der Grundlage patientenrelevanter Endpunkte (Glaeske 2011).

Die Realität

Dennoch werden im Gesundheitswesen wegweisende Entscheidungen häufig aufgrund von Surrogaten getroffen. Die Studie von Ujeyl et al (2012) zur Verfügbarkeit von Daten bei der Markteinführung von Arzneimitteln zeigt, dass bei der Zulassung von Medikamenten nur in 46% der Fälle ein primärer patientenrelevanter Endpunkt erhoben worden war.

Der gemeinsame Bundesausschuss (G-BA) beschließt, welche Interventionen von der GKV übernommen werden. Bei diesen Entscheidungen ist auch er den Prinzipien der evidenzbasierten Medizin verpflichtet, wonach neue Methoden nur anhand von „patientenrelevanten Auswirkungen“ zu bewerten sind (G-BA Methodenbewertung). Faktisch wurde dieses Prinzip jedoch immer wieder unterlaufen: Mit Beschluss vom 15.07.2007 wird das Hautkrebsscreening ab 35 als GKV-Leistung beschlossen (BAnz. Nr. 37, S. 871) vom 06.03.2008). Aus den tragenden Gründen zum Beschluss geht eindeutig hervor, dass keine Evidenz für patientenrelevante Endpunkte wie die Reduktion der melanombedingten Mortalität vorliegt. Begründet wird der Beschluss mit dem Surrogat „Vorverlegung des Diagnosezeitpunkts“ und der Möglichkeit, Frühstadien des Melanoms kurativ zu behandeln (G-BA 2007). Früherkennung ist ein Angebot für die gesunde Gesamtbevölkerung, das von mehr Menschen in Anspruch genommen werden kann als eine Maßnahme bei einer bestimmten Indikation. Der G-BA hat hier eine Entscheidung von erheblichen finanziellen Konsequenzen und unklarem Nutzen für die Patienten aufgrund von Surrogatendpunkten getroffen.

Mit der Entscheidung des G-BA sowie den Vorgaben von Zertifizierungskommissionen, Mindestmengen für bestimmte Prozeduren vorzuschreiben, um Ergebnisqualität zu sichern, halten Surrogate auch Einzug in die Qualitätssicherung. Es liegen keine Daten vor, die einen kausalen Zusammenhang zwischen Menge und Ergebnisqualität belegen (Weißbach/Schaefer 2010). Wenn dies auch nach Auffassung des Bundessozialgerichts (Urteil vom 12.09 und 18.12.1012) nicht notwendig und die Festsetzung von Mindestmengen aus juristischer Sicht grundsätzlich zulässig sei, so bleibt die Menge ein Surrogat für Ergebnisqualität, das nach der gängigen Definition (vergl. Einleitung) kein valider Parameter ist, da der Nachweis des Kausalzusammenhangs fehlt. Dies ist angesichts der Anreize, die vor allem durch die Zertifizierungsvorgaben entstehen, höchst bedenklich. Nicht nur der „Transplantationsskandal 2012/2013“ hat gezeigt, welche Auswirkungen der unkritische Umgang mit einem invaliden Surrogat zeitigen kann.

Konsequenzen

Wie eingangs gezeigt, ist theoretisch gut fundiert, dass Surrogate mit Unsicherheit behaftet sind und empirisch belegt, dass die Anwendung in vielen Fällen zu Trugschlüssen bezüglich patientenrelevanter Endpunkte geführt hat. In Studien werden sie oft eingesetzt, weil erst durch die Wahl der (Surrogat-) Zielparameter ein signifikanter (positiver) Effekt der Intervention belegt werden kann. Die Folge: Erkenntnisse über die beste – oder wenigstens die angemessene – Versorgung werden unter anderem durch die Anwendung von Surrogaten verzerrt, und Patienten erhalten in der Folge eine weniger nützliche oder eine schädlichere Intervention. Diese für die Versorgung und das Wohl der Patienten verheerende Praxis darf nicht akzeptiert, sondern muss beendet werden. An Konsequenzen sind zu ziehen: Ausbau der wissenschaftlichen Fundierung von Entscheidungen im Gesundheitswesen und die konsequente Beteiligung von Patientinnen und Patienten.

Die „richtigen“ Studien

Die erste Konsequenz aus den Unsicherheiten ist daher die Forderung nach den „richtigen“ Studien. Der Gesetzgeber hat in der Arzneimittel-Nutzenbewertungs-Verordnung vom 28. Dezember 2010 den Nutzen von Arzneimitteln unmissverständlich nur in Bezug auf patientenrelevante Endpunkte definiert. Es scheint daher logisch und geboten, diese Definition auch bei der Zulassung heranzuziehen. Mithin wäre eindeutig zu fordern, dass die Parameter, die lt. Gesetz zur Bewertung des Nutzens

gelten sollen, auch für die Zulassung verpflichtend sind. Zulassungsstudien müssen primär auf die patientenrelevanten Endpunkte Morbidität, Mortalität und Lebensqualität testen. Andere sollten nicht akzeptiert werden. Und analog der Initiative des International Committee of Medical Journal Editors, nur noch prospektiv registrierte Studien zur Veröffentlichung anzunehmen, um Publikationsbias zu vermeiden (ICMJE recommendations, 2010), könnte eine Veröffentlichung von Studien ebenso an die Verpflichtung zum Bericht patientenrelevanter Endpunkte geknüpft werden. Auch das CONSORT-Statement (the CONSORT Group 2010), das Standards für den Ergebnisbericht klinischer Studien formuliert, sollte deren Berücksichtigung eindeutig fordern und definieren, wann Surrogate zulässig sein können. Vielleicht werden vom US-amerikanischen Patient-Centered Outcomes Research Institute (www.pcori.org), in das viele hochrangige Repräsentanten des amerikanischen Gesundheitssystems eingebunden sind, wichtige Impulse ausgehen, patientenrelevante Endpunkte als gute Praxis für alle klinischen Studien zu verankern.

Patientenbeteiligung

Eine zweite, entscheidende Konsequenz ist die Forderung nach umfassender Patientenbeteiligung auf allen Ebenen des Gesundheitssystems - auch bei der Planung und Bewertung von Studien zur Gesundheitsforschung. Patienten sind unmittelbar von möglichen Schäden falscher Systementscheidungen betroffen. Je stärker ihre Perspektive in den Entscheidungsprozess eingebunden wird, desto größer ist die Chance, dass nach Kriterien entschieden wird, die für sie tatsächlich relevant sind. Voraussetzung dafür ist jedoch, dass Patientenbeteiligung ernstgenommen, nach einer definierten Methodik (z.B. einem transparenten Benennungsverfahren) umgesetzt und die Vertreter entsprechend geschult werden (Berger 2010). Dabei ist zunächst zu unterscheiden zwischen der Beteiligung auf Systemebene - also in entscheidungsrelevanten Gremien, bei Wissensgenerierung und –transfer - einerseits und auf der individuellen Ebene - das heißt, bei der partnerschaftlichen Entscheidungsfindung - andererseits.

Systemebene

Bei Entscheidungen auf der gesundheitspolitischen Ebene sind Patienten heute nicht nur beteiligt, die Beteiligung ist sogar gesetzlich verpflichtend (SGB V §140f, g) und das Prozedere in der „Patientenbeteiligungsverordnung“ (PatBeteiligungsV, BGBl. I S. 2753) gesetzlich geregelt. Bestimmte Patientenorganisationen haben das Recht, Vertreter in den G-BA zu entsenden, die dort bei allen Entscheidungen ein Anhörungs-, aber kein Stimmrecht haben.

Beim Wissenstransfer, zum Beispiel bei der Erstellung von medizinischen Leitlinien oder Patienteninformationen, ist die Einbeziehung von Patienten nicht verpflichtend, wird aber als Qualitätskriterium definiert – national wie international (DELBI, AGREE, GRADE, G-I-N Public Toolkit, GPGI). In Leitliniengruppen besitzen sie volles Stimmrecht (ÄZQ 2008), derzeit aber sind sie deutschlandweit nicht einmal bei der Hälfte aller hochwertigen Leitlinien (Entwicklungsstufe/S 3) beteiligt (Schaefer 2010).

Bei der Erarbeitung von HTA-Berichten und systematischen Reviews werden Patienten bislang nur sporadisch einbezogen (Kreis 2012, Facey 2010), völliger Nachholbedarf besteht bei der Konzeption und Durchführung klinischer Studien. Obwohl die Forderung nach Patientenbeteiligung gerade in diesem Bereich vehement erhoben wird (European Science Foundation 2012), gibt es bislang nur wenige entsprechende Initiativen (z.B. die James Lind Alliance in Großbritannien). Dabei wäre vor allem hier die Patientenperspektive unerlässlich, um schon in der Konzeption klinischer Studien die maßgebliche Berücksichtigung patientenrelevanter Endpunkte durchzusetzen und invalide Surrogate weitestgehend zu vermeiden.

Zusammenfassend lässt sich ein Konsens darüber feststellen, dass die Beteiligung von Patienten auf der Systemebene wünschenswert ist, weil sie die Berücksichtigung patientenrelevanter Endpunkte und so die Relevanz von systempolitischen Entscheidungen sichert. Dennoch wird sie selten praktiziert.

Partizipative Entscheidungsfindung (PEF)

Auf der Ebene der individuellen Arzt-Patienten-Beziehung gilt heute allgemein als selbstverständlich, dass Patienten in die Entscheidungsfindung einbezogen werden (Härter et al 2011). Idealerweise können sie dann Nutzen und Schaden von therapeutischen oder diagnostischen Interventionen gegen Bedürfnisse, Wertvorstellungen und Lebenssituation abwägen, um eine für sie „richtige“ Entscheidung zu treffen. Dazu müssen sie die verfügbaren und verlässlichen Daten zu Nutzen und Schaden der Entscheidungsalternativen kennen und verstehen (Klemperer et al 2009). Sie brauchen also verständliche, systematisch recherchierte und hochwertige Informationen. Diese Informationen müssen Endpunkte berichten, die in ihrer Entscheidungssituation von Bedeutung sind. Ein aufgrund von invaliden Surrogaten suggerierter Nutzen hingegen kann zu einer gesundheitsgefährdenden Entscheidung führen. So wird z. B. bei der Früherkennung auf Eierstockkrebs ein verbessertes Fünf-Jahres-Überleben nach transvaginalem Ultraschall (TVUS) als Nutzenbeleg kommuniziert (van Nagell 2011). Das verbesserte Fünf-Jahres-Überleben ergibt sich aber einzig aus der Vorverlegung des Diagnosezeitpunkts und ist an sich kein valider Parameter für den Nutzen der Früherkennungsuntersuchung (Wegwarth/Gigerenzer 2011). Tatsächlich wird infolge des TVUS kein einziger Tod verhindert, aber es werden in Folge von Überdiagnosen viele zusätzliche Ovarektomien vorgenommen, die Patienteninnen schädigen, ohne den Verlauf der Erkrankung zu beeinflussen (Buys 2011).

Das neue Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten (Patientenrechtegesetz - PatRG 2013) betont die Bedeutung von Aufklärung und Information für den Patienten. Diese Akzentuierung wird zum Beispiel an der dezidierten Unterscheidung zwischen Sicherungsaufklärung (BGB § 630c) zur Beteiligung und Gefahrabwendung sowie der Selbstbestimmungsaufklärung (§ 630e) zur Verlaufs-, Risiko- und Alternativenaufklärung deutlich.

Bislang sind für den Arzt jedoch nur grobe Angaben zu Nutzen und Nebenwirkungen rechtlich verpflichtend, beide müssen nicht beziffert werden: Nach der allgemeinen Rechtsprechung hat der Arzt den Patienten nur „im Großen und Ganzen“ über Chancen und Risiken einer Behandlung aufzuklären (BGHZ 90, 103ff). Das Sondergutachten des Sachverständigenrates 2012 betont die Notwendigkeit von hochwertiger Patienteninformation ebenfalls und fordert eine zentrale Einrichtung für diese Aufgabe. Doch die Grundlagen der Aufklärung (z.B. valide patientenrelevante Endpunkte, hochwertige Studien) werden weder in der Rechtsprechung noch im Patientenrechtegesetz oder im aktuellen Gutachten des Sachverständigenrates zur Begutachtung der Entwicklung im Gesundheitswesen (2012) adressiert.

Auf der Systemebene können beteiligte Patienten darauf dringen, dass Surrogate nicht berichtet oder auf ihrer Grundlage keine Handlungsempfehlungen ausgesprochen bzw. Systementscheidungen getroffen werden. Auf der individuellen Ebene können Patienten nur darauf achten, ob für die Optionen ihrer aktuell anstehende Entscheidung Nutzen und Schaden anhand von harten Endpunkten untersucht und ob diese verständlich berichtet wurden. Und sie können entscheiden, ob diese

Literatur

The AGREE Next Steps Consortium (2009): Appraisal of guidelines for research & evaluation II instrument. In: http://www.agreetrust.org/index.aspx?o=1397 (abgerufen am 18.01.2013)

ÄZQ (Hrsg.) (2008): Handbuch Patientenbeteiligung. ÄZQ Schriftenreihe Bd. 33. In: http://www.aezq.de/mdb/edocs/pdf/schriftenreihe/schriftenreihe33.pdf (abgerufen am 13.01.2013)

Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) und Ärztliches Zentrum für Qualität in der Medizin (Hrsg.) (2008): Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI)Fassung 2005/2006 + Domäne 8 (2008). Berlin 2008. In: www.delbi.de (abgerufen am 18.1.2013)

Berger, B./Steckelberg, A./Meyer, G./Kasper, J./Mühlhauser I (2010): Training of patient and consumer representatives in the basic competencies of evidence-based medicine: a feasibility study. BMC Medical Education, 10:16

Booth, CM./Cescon, DW./Wang, L./Tannock, IF./Krzyzanowska, MK. (2008). Evolution of the randomized controlled trial in oncology over three decades. J Clin Oncol 26:5458-5464

BT-D (2010): Bundestagsdrucksache 17/4243 vom 10.12.2010: Unterrichtung durch die Bundesregierung: Rahmenprogramm Gesundheitsforschung der Bundesregierung.

Bundessozialgericht: Urteile zu Mindestmengen Knie TEP (12.09.2012) und Frühgeborene (18.12.2012)

Buys, SS. et al. (2011): Effect of screening on ovarian cancer mortality: the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial. JAMA. 2011 Jun 8;305(22):2295-303

Buyse, M. (2009): Use of meta-analysis for the validation of surrogate endpoints and biomarker in cancer trials. The Cancer Journal, 15:5, 421-5

CONSORT Group (2010): The CONSORT Statement. In: http://www.consort-statement.org/consort-statement/ (abgerufen am 13.01.2013)

Deutscher Ethikrat (Hrsg.) (2012): Dokumentation der Jahrestagung 2012, Personalisierte Medizin - der Patient als Nutznießer oder Opfer?, Berlin, im Druck.

Dubben, HH. (2009): Trials of Prostate cancer screening are not worthwhile. Lancet Oncol. 2009 Mar;10(3):294-8

European Medicines Agency completes its review of Avastin used in breast cancer. Press release 2012/16/12. In: http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2010/12/news_detail_001166.jsp&mid=WC0b01ac058004d5c1 (abgerufen am 13.01.2013)

European Science Foundation 2012: Science Policy Briefing 45. Implementation of Medical Research in Clinical Practice. In: http://www.esf.org/index.php?eID=tx_ccdamdl_file&p[file]=41634&p[dl]=1&p[pid]=4052&p[site]=European%20Science%20Foundation&p[t]=1358675754&hash=5a34d67c59fd45fe868a0b278f6c5d66&l=en (abgerufen am 13.01.2012)

Facey, K./Boivin, A./Gracia, J./Hansen, HP./Lo, Scalzo, A./Mossman, J./Single, A. (2010): Patients' perspectives in health technology assessment: a route to robust evidence and fair deliberation. Int J Technol Assess Healt: h Care. 2010 Jul;26(3):334-40

Fleming, TR./Powers, JH. (2012): Biomarkers and surrogate endpoints in clinical trials. Stat Med. 2012 Nov 10;31(25):2973-84

G-BA. Methodenbewertung. In: http://www.g-ba.de/institution/auftrag/arbeitsweise/methodenbewertung/ (abgerufen am 12.01.2013)

G-BA: Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungs-Richtlinien: Hautkrebs-Screening vom 15. November 2007. In: http://www.g-ba.de/downloads/40-268-482/2007-11-15-KFU-Hautkrebsscreening_TrG.pdf (abgerufen am 12.01.2013)

G-I-N PUBLIC. Toolkit on Public and Patient Involvement in Guidelines (2012). In: http://www.g-i-n.net/activities/gin-public/toolkit (abgerufen am 13.01.2013)

Glaeske, G. (2012): Surrogatendpunkte - irgendwo zwischen unverzichtbar und indiskutabel. ZEFQ 106, 161-167 (dazu Vortrag auf dem IQWIG Herbst Symposium 25.11.2011. Im Internet:
https://www.iqwig.de/download/11-11-25_Glaeske_Surrogatendpunkte.pdf (abgerufen am 15.01.2013)

Hüsing B et al (2008): Individualisierte Medizin und Gesundheitssystem. Arbeitsbericht Nr. 126. Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag. Berlin. In: http://www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Arbeitsbericht-ab126.pdf (abgerufen am 19.01.2013)

International Committee of Medical Journal Editors (2010): Obligtion to register clinical trials. In: http://www.icmje.org/publishing_10register.html (abgerufen am 12.01.12013)

Klemperer, D./Lang, B./Koch, K. et al. (2010): Gute Praxis Gesundheitsinformation. Z. Evid. Fortbild. Qual. Gesundh. wesen 104 (2010) 66–68

Kreis, J./Puhan, MA./Schünemann, HJ./Dickersin ,K. (2012): Consumer involvement in systematic reviews of comparative effectiveness research. Health Expect. 2012 Mar 6

Krogsbøll, LT./Jørgensen, KJ./Grønhøj Larsen, C./Gøtzsche, PC. (2012): General health checks in adults for reducing morbidity and mortality from disease (Review). The Cochrane Library 2012 (10)

Kunz, R./Burnand, B./Schünemann HJ. (2008): Das GRADE-System. Der Internist 2008(6):673-80

Mangiapane, S./Garrido, MV (2009): Surrogatendpunkte als Parameter der Nutzenbewertung. Schriftenreihe Health Technology Assessment, Bd. 91. In: http://portal.dimdi.de/de/hta/hta_berichte/hta250_bericht_de.pdf (abgerufen am 15.01.2013)

Mühlhauser, I./Berger, M. (1996): Surrogat-Marker: Trugschlüsse. Dtsch Ärztbl 93:A3280-3

Mühlhauser, I./Meyer, G. (2006): Evidence Based Medicine: Widersprüche zwischen Surrogatergebnissen und klinischen Endpunkten. Psychother Psych Med;56:193-201

Mühlhauser, I./Müller, H. (2009): Patientenrelevante Endpunkte und patient-reported outcomes in klinischer Forschung und medizinischer Praxis. In: Klusen, N/Fließgarten, A/Nebling, T (Hrsg.): Informiert und selbstbestimmt. Der mündige Bürger als mündiger Patient. Beiträge zum Gesundheitsmanagement, Bd. 24, 34-65

Müller, H. (2012): Chancen und Risiken der „individualisierten Medizin“ für das Gesundheitssystem. Welt der Krankenversicherung 2: 40-46

Patientenbeteiligungsverordnung vom 19. Dezember 2003 (BGBl. I S. 2753), die durch Artikel 457 der Verordnung vom 31. Oktober 2006 (BGBl. I S. 2407) geändert worden ist. In: http://www.gesetze-im-internet.de/patbeteiligungsv/BJNR275300003.html. Letzter Aufruf: 13.01.2013; 10:34

PatRG (2013): Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten. Bundesrat Drucksache 7/13

Referentenentwurf des Bundesministeriums für Gesundheit: Entwurf eines Gesetzes zur Stärkung der Gesundheitsförderung und der Prävention im Gesundheitswesen (Gesundheitsförderungs- und Präventionsstärkungsgesetz), Stand 11.01.2013

Referentenentwurf des Bundesministeriums für Gesundheit. Entwurf eines Gesetzes zur Umsetzung der Empfehlungen des Nationalen Krebsplans (07/2012) In: http://www.tumorzentren.de/tl_files/dokumente/Nationaler%20Krebsplan/120702_Referentenentwurf_Krebsplan-Umsetzungsgesetz.pdf (abgerufen am 16.01.2013)

Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen (2012): Wettbewerb an den Schnittstellen der Gesundheitsversorgung (Sondergutachten 2012). In: http://dip21.bundestag.de/dip21/btd/17/103/1710323.pdf (abgerufen am 18.01.2013)

Schaefer, C./Nothacker, M./Langer, T./Conrad, S./Fishman, L./Meyerrose B. et al. (2010):Patient involvement in Germany: How and when?. Otolaryngol Head Neck Surg 2010;143(1 Supplement 1):50

Schaefer, C./Dubben, HH./Weißbach L. (2011): Wer sorgt hier vor? Der Onkologe 2011; 17: 220-34

Schumpelik V, Vogel B (Hrsg.) (2011): Medizin nach Maß. Individualisierte Medizin - Wunsch und Wirklichkeit.

Sozialgesetzbuch V. Buch. § 140f und 140g. In: http://www.sozialgesetzbuch-sgb.de/sgbv/140f.html (abgerufen am 13.01.2013)

Ujeyl, M. et al (2012): Neue Arzneimittel: Verfügbarkeit von Daten zum therapeutischen Stellenwert bei Markteinführung. Dtsch Arztebl Int 2012; 109(7): 117-23; DOI: 10.3238/arztebl.2012.0117

US Food and Drug Administration (FDA). Avastin Information. In: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm193900.htm (abgerufen am 13.01.2013)

van Nagell, JR. et al. (2011): Long-term survival of women with epithelial ovarian cancer detected by ultrasonographic screening. Obstet Gynecol. 2011 Dec;118(6):1212-21

Verordnung über die Nutzenbewertung von Arzneimitteln, Bundesgesetzblatt Jahrgang 2010, Teil I Nr. 68, ausgegeben zu Bonn, 31.Dezember 2010. In: http://www.bgbl.de/Xaver/start.xav?startbk=Bundesanzeiger_BGBl&bk=Bundesanzeiger_BGBl&start=//*[@attr_id=%27bgbl110s2324.pdf%27] (abgerufen am 12.01.2013)

Wagner, A./Thomssen, S./Haerting, J./Unverzagt, S. (2012): Vascular-endothelial-growth-factor (VEGF) targeting therapies for endocrine refractory or resistant metastatic breast cancer. The Cochrane Library, (2012/7/11)

Wegwarth, O./Gigerenzer, G. (2011): Risikokommunikation: Sterblichkeitsstatistik als valides Maß Dtsch Arztebl 2011; 108(14): A-760 /B-624 / C-624

Weißbach, L./Schaefer, C. (2010): Kontra Mindestmengen: Mindestmengen in der onkologischen Chirurgie sind kein verlässliches Instrument zur Qualitätssicherung! best practice onkologie (2/2010); 5 (1):16-23

Verordnung über die Nutzenbewertung von Arzneimitteln, Bundesgesetzblatt Jahrgang 2010, Teil I Nr. 68, ausgegeben zu Bonn, 31.Dezember 2010; §2, Abs.3)

Endpunkte in ihrer persönlichen Situation bzw. für sie selbst ebenfalls relevant sind.

Aktuelle Herausforderungen

Die Umsetzung der beiden vorangehenden Forderungen nach aussagekräftigen Studien und mehr Patientenbeteiligung kann dazu beitragen, dass viele System- und Behandlungsentscheidungen nicht mehr auf der Grundlage von Surrogaten getroffen werden und damit für Fehlschlüsse anfällig sind. Ganz konkret erscheinen einige aktuelle Entwicklungen im Gesundheitssystem in einem neuen Licht, wenn man sie ausschließlich auf ihren Nutzen hinsichtlich patientenrelevanter Endpunkte analysiert und invalide Surrogate bei der Nutzenbewertung für unzulässig erklärt.

Personalisierte Medizin

Die sogenannte „Personalisierte Medizin“ ist ein Schwerpunkt der Innovationspolitik und der Gesundheitsforschung (BT-D 2010, Rahmenprogramm Gesundheitsforschung der Bundesregierung, Hüsing et al. 2008). Mit Macht drängt sie in die gesundheitspolitische Diskussion (Deutscher Ethikrat 2012; Schumpelik, Vogel 2011; zu der Relevanz des Konzeptes für die GKV Müller 2012). Der Begriff vermischt zwei Konzepte:

Zum einen verbindet sich damit das Versprechen, jedem Patienten die für seine biologische und genetische Disposition maßgeschneiderte Behandlung anbieten zu können: Für bestimmte genetische oder biologische Merkmale werden spezifische Wirkstoffe entwickelt. Indem das Wirkspektrum der Pharmaka auf eine enge Patientengruppe eingegrenzt wird, werden aus einer häufigen Erkrankung lauter einzelne, seltenere Erkrankungen (Orphanisierung). Voraussetzungen für solide Evidenzbasierung werden dadurch erschwert. Gleichzeitig drängen die hochpreisigen Medikamente mit Druck auf den Markt, häufig sollen die Substanzen im beschleunigten Verfahren und aufgrund von Zwischenergebnissen und Surrogatendpunkten zugelassen werden.

Beispielhaft lässt sich dies an einem „Blockbuster“ der Personalisierten Medizin ablesen: Bevacizumab, ein Angiogenesehemmer der Firma Roche, wurde 2008 im beschleunigten Verfahren für die Behandlung bei metastasiertem Brustkrebs zugelassen. Grundlage für die Zulassung war eine signifikante Reduktion des progressionsfreien Überlebens in einer Zwischenauswertung, die von den Zulassungsbehörden als Surrogat für die Reduktion des Gesamtüberlebens akzeptiert wurde. Im November 2011 hatte die amerikanische FDA dem Wirkstoff die Zulassung für diese Indikation jedoch wieder entzogen: Die Analyse der nun vorliegenden Daten zeigte trotz längerem progressionsfreiem Intervall keine Verbesserung des Gesamtüberlebens (FDA 18/11/2011). Der aktuelle Cochrane Review kommt zum selben Ergebnis (Wagner 2012). Die Europäische Arzneimittelbehörde EMA sieht dennoch keine Veranlassung, die Zulassung zurückzuziehen. Für sie reicht das progressionsfreie Überleben als Surrogat völlig aus (EMA 2012). Welche enormen finanziellen Auswirkungen die FDA-Entscheidung hatte, lässt sich aus den Veröffentlichungen in führenden Wirtschaftsmagazinen ablesen: Das umsatzstärkste Medikament der Firma Roche brach ein, der Aktienkurs ebenfalls. Inzwischen wird mit einer höchst zweifelhaften „pay-for-performance“-Strategie der Umsatz in die Höhe getrieben: Wenn Bevacizumab nicht wirkt, erstattet Roche dem Krankenhaus die Behandlungskosten und schafft so einen Anreiz zum Schaden der Patientinnen. Was für die Patientinnen relevant ist, zählt jedoch nicht: Längst ist das, was als Versprechen daherkam, ein aggressiv expandierender Markt. Surrogate sind ein Mittel von mehreren, den in der Nutzen-Schadensbilanz oft zwiespältigen Medikamenten einen Marktzugang zu verschaffen.

Zum Zweiten wird mit der „Personalisierten Medizin“ ein Prädiktionskonzept verbunden, nach dem Erkrankungs-Risiken und Behandlungs-Chancen einzelner Menschen ausschließlich anhand von deren biologischen Merkmalen (z.B. genetischer, zytologischer, morphologischer Informationen) eingeschätzt werden. Die medizinische Behandlung wie auch die Unterlassung der Behandlung werden durch operationalisierbare biologische Eigenschaften des Individuums determiniert und legitimiert. Bei dieser Form der Prädiktion wird der Surrogatparameter zum Prinzip erhoben: Bio- und genetische Marker sind Surrogate im klassischen Sinne: Es gibt zwar einen retrospektiv nachgewiesenen Zusammenhang zwischen dem Marker und der Häufigkeit bestimmter Erkrankungen bzw. einem bestimmten Krankheitsverlauf. Valide sind diese Surrogate jedoch im Sinne der Methodik nicht: Es gibt keine Nachweise zur Kausalität, d.h. dass mit der Beeinflussung des Markers auch Morbidität, Mortalität oder Lebensqualität beeinflusst werden. Dies ist nur mit randomisierten kontrollierten Studien möglich, die auf diese harten, patientenrelevanten Endpunkte testen (Buyse 2009, Fleming & Powers 2012). Je weiter der Diagnosezeitpunkt nach vorn verlegt und Prä-Erkrankungen diagnostiziert werden, desto schwerer ist es zudem, in einem randomisierten Studiendesign einen Zusammenhang zwischen den Surrogaten und patientenrelevanten Outcomes nachzuweisen: Die möglichen Varianzen in der weiteren Diagnostik und Behandlung werden immer größer und die Ergebnisse damit immer weniger vergleichbar. Prädiktion im Sinne der „Personalisierten Medizin“ basiert daher ausschließlich auf nicht validierten Surrogaten. Ob diese Prädiktion einen tatsächlichen Nutzen in Bezug auf Krankheitslast, Sterblichkeit oder Lebensqualität hat, wird vermutlich nicht nachzuweisen sein. Die Schäden hingegen liegen auf der Hand: Statt Krankheiten werden Erkrankungsrisiken diagnostiziert, deren tatsächliches Eintreten jedoch ungewiss ist und deren Erkenntniswert für den Patienten oft höchst fraglich ist. So werden viele Menschen unnötig zu Patienten (und ggf. unnötig behandelt).

Prävention

Eine ähnliche Problematik entspannt sich auch im Bereich der Prävention. Es hat sich zunehmend die Überzeugung durchgesetzt, dass es besser sei, Krankheiten zu vermeiden, statt sie später mit hohem Aufwand zu therapieren. Dem entspringen viele Initiativen zur Prävention, etwa die regelmäßigen Gesundheits-„Check-Ups“, Kampagnen für einen gesunden Lebensstil, Maßnahmen zur Früherkennung, wie sie zum Beispiel mit dem Krebsplanumsetzungsgesetz (Referentenentwurf 07/2012) durchgesetzt werden sollen. Der Referentenentwurf für ein Gesundheitsförderungs- und Präventionsstärkungsgesetz (Referentenentwurf „Präventionsgesetz“ 2013“, Stand 11.01.2013) formuliert deutlich das Ziel, durch die Verstärkung präventiver Maßnahmen Krankheit zu vermeiden, die Gesundheit der Menschen zu fördern und Krankheits- und Krankheitsfolgekosten zu reduzieren. Als Maßnahme zur Zielerreichung nennt der Referentenentwurf auch die „Sicherstellung der Qualität und Förderung der Wirksamkeit von Prävention und Gesundheitsförderung“ - ohne allerdings zu diesem Punkt dezidierte Regelungen im Gesetz vorzusehen und ohne die Evidenzbasierung präventiver Maßnahmenempfehlungen überhaupt zu stärken. Dies ist insofern ungewöhnlich, als üblicherweise ein Gesetz zu den mit dem Gesetz angestrebten Lösungen beiträgt und gerade der evidenzbasierte Wirksamkeitsnachweis und die Sicherstellung der Qualität von präventiven Maßnahmen eine Herausforderung darstellen.

Die Präventionsziele sind ehrenwert, allerdings beruhen die bisherigen Annahmen zum Nutzen der Prävention größtenteils auf Surrogaten. Zu patientenrelevanten Endpunkten gibt es nur wenige Daten, die jedoch nachdenklich machen: Ein aktueller Cochrane Review zeigt, dass so genannte allgemeine Gesundheitschecks (Check-Ups) keinen Effekt auf Morbidität und Mortalität der Bevölkerung haben, sondern al-

lein die Gesamtkosten für Diagnostik und Behandlung signifikant erhöhen (Krogsboll 2012). Das steht der Forderung des Präventionsgesetzes, Check-Ups häufiger und früher anzubieten, diametral entgegen.

Für die Krebsfrüherkennung wird inzwischen diskutiert, ob der Nachweis der Reduktion tumorspezifischer Mortalität als Nutzen ausreicht (Schaefer et al 2011). Eine durch Krebsfrüherkennung bedingte Reduktion der Gesamtmortalität nachzuweisen ist aus statistischen Gründen nahezu unmöglich, dafür wären Studien mit Millionen von Teilnehmern notwendig, die vermutlich rein logistisch nicht durchführbar sind (Dubben 2009). Auch ein Effekt des Screenings auf die Lebensqualität wird nicht valide erhebbar sein: Ein Problem, das bei langfristigen Studien zur Lebensqualität auftritt, ist das Phänomen des „Response-Shifts“: Es kommt generell bei jedem Menschen zur Adaptation an neue Situationen, die ihn bestimmte Ereignisse anders beurteilen lassen. (Aufgefallen war dieses Phänomen in der Frühphase der EORTC QoL group, da auch bei Messung der Lebensqualität an mehr als 10.000 Krebspatienten kaum therapieabhängige oder stadienspezifische Veränderungen nachweisbar waren.) Die gesetzlich geforderten patientenrelevanten Endpunkte werden also für Früherkennungsmaßnahmen prinzipiell nicht überprüfbar sein.

Zu guter Letzt muss auch gefragt werden, ob die Annahmen zum Nutzen der Primärprävention den Anforderungen an valide Endpunkte standhalten. Prävention fokussiert auf eine generelle Gesunderhaltung. Bislang werden aber nur bestimmte indikationsbezogene Endpunkte untersucht, beim Rauchen etwa die Häufigkeit von bösartigen Neubildungen, Atemwegserkrankungen und kardiovaskulären Ereignissen. Diese werden als Surrogate für allgemeine Gesundheit akzeptiert. Für keinen dieser Endpunkte wurde bislang nach der oben dargestellten Methodik im randomisierten kontrollierten Design nachgewiesen, dass deren Beeinflussung kausal die (Gesamt!)Mortalität, Morbidität oder Lebensqualität verbessert. Letztlich beruhen damit auch unsere Annahmen zum Nutzen der Primärprävention auf nicht validierten Surrogaten. Das muss nicht gegen Prävention sprechen. Nur sollte es dazu anhalten, den Nutzen von Prävention nicht zu überschätzen und Gesunderhaltung nicht als Pflicht anzusehen. „Um dement zu werden, muss ich erst einmal den Herzinfarkt überleben“, hat ein Gesundheitspolitiker einmal formuliert. Welche der beiden Prognosen für einen Menschen relevanter ist, das kann nur er vor dem Hintergrund seiner persönlichen Erfahrungen und Wertvorstellungen entscheiden.

Fazit

Surrogate können nur, wenn sie validiert wurden, als verlässliche Endpunkte gelten. Die für diese Validierung notwendige Kausalität zu relevanten klinischen Endpunkten konnte bislang für die wenigsten Surrogate nachgewiesen werden. Im Gegenteil wurde bei Interventionen, die die erwünschten Veränderungen in den Surrogatparametern erzielten, später ein abweichender, schädlicher Effekt auf die klinischen Endpunkte beobachtet. Studien, die ausschließlich Surrogate berichten, liefern daher keine verlässlichen Ergebnisse. Dennoch werden in der Versorgungsrealität immer wieder wichtige Entscheidungen aufgrund von Surrogaten getroffen. Es ist notwendig, die sich daraus ableitenden Konsequenzen endlich zu ziehen: die Forderung nach Studien mit klinisch relevanten Endpunkten endlich in allen Regelwerken zu verankern sowie Patienten konsequent auf allen Entscheidungsebenen der Gesundheitsversorgung zu beteiligen. Vor allem ist es dringend notwendig, die anstehenden Entwicklungen und Entscheidungen im Gesundheitssystem daraufhin zu prüfen, ob die ihnen zugrundeliegenden Annahmen auf validen patientenrelevanten Endpunkten beruhen. <<

Surrogates and patient-centered outcomes: Fallacies and consequences

Surrogates are indicators representing a preliminary stage of a clinical endpoint non-perceivable for patients. They often are used in clinical trials, since they allow for less complex study designs than clinical endpoints. Surrogates have to be validated to provide reliable data. To this end, RCTs must show that any therapeutic effect on the surrogate causally affects the clinical endpoint. Nevertheless, most surrogates that are currently in use have not been validated. Consistently, this has led to fallacies, that is: observed positive effects on the surrogate entailed no or adverse effects on the clinical endpoint in question. Treatment or general health care decisions therefore should not be based on surrogate endpoints, but that is what happens most often in clinical practice. As a consequence of this unsustainable situation, that affects patient safety, the following claims have to be pushed: reporting clinical endpoints should be mandatory for drug admission and a standard for any clinical trial to provide data on patient-centered outcomes; patients and consumers should be involved at all stages and levels of health care decisions to assure any clinical or health care decision only is made on the basis of patient-centered outcomes. And, most urgently, it has to be analyzed if underlying presumptions of current developments and decisions in health care are based on valid endpoints.

Keywords

surrogates, surrogate endpoints, patient-centered outcomes, benefit, shared decision making, personalized medicine, patient rights, prevention

Hardy Müller

Nach den Studium der Anthropologie, Soziologie und Psychologie wissenschaftlicher Mitarbeiter in einem Forschungsschwerpunkt „Soziale Determinanten regionaler Mortalitätsunterschiede“. Seit 1993 in der gesetzlichen Krankenversicherung tätig. Arbeitsschwerpunkte sind Analysen von Routinedaten, Evaluation medizinischer Leistungen, evidenzbasierte Patienteninformation und Patientensicherheit. Ehrenamtlicher tätig als Geschäftsführer des Aktionsbündnis Patientensicherheit e.V. Kontakt: hardy.mueller@wineg.de



Corinna Schaefer

Studium der Romanistik, Musik und Theaterwissenschaft, künstlerische und journalistische Tätigkeit, seit 2006 am Ärztliche Zentrum für Qualität in der Medizin, seit 2010 Leitung des Bereichs Patienteninformation und –beteiligung am Ärztlichen Zentrum für Qualität in der Medizin; Chair der Public and Patient Involvement Working Group des Guidelines International Network. Kontakt: schaefer@azq.de



Prof. Dr. Axel C. Mühlbacher
Susanne Bethge, M.Sc.
Christin Juhnke, M.A.

„Patientenrelevante Endpunkte“ - Bewertung aus der Perspektive der Patienten?

Die meisten Entscheidungsträger, einschließlich IQWiG, berücksichtigen verschiedene patientenrelevante Endpunkte, um die Überlegenheit einer Therapie gegenüber einem Placebo oder einen Komparator zu beweisen. Es stellt sich die Frage: Wie können widersprüchliche Ergebnisse in verschiedenen Endpunkten interpretiert werden? Das Problem in der Nutzenbewertung mit mehr als einem Endpunkt ist die Gewichtung und die darauf aufbauende Berechnung einer Kennzahl für den Gesamtnutzen. Ohne die Aggregation von mehreren Endpunkten zu einem Score/ Index könnte die Ableitung von Empfehlungen oder regulatorischen Entscheidung fraglich sein.

Zusammenfassung

Der dokumentierte Patientennutzen sollte ausschlaggebend für die Prioritätensetzung sein. Die zugrundeliegenden Entscheidungsprozesse müssen dafür Sorge tragen, dass die Auswahl aus einer Vielzahl potentieller Gesundheitsgüter und -dienstleistungen mit den Bedürfnissen, Erwartungen und Prioritäten der Versicherten und Patienten im Einklang stehen. Die gemeinsame Basis des klinischen und des ökonomischen Entscheidungskalküls ist der Nutzen für die Patienten. Patientenpräferenzen sollten deshalb auf allen Ebenen in die Entscheidungsprozesse einbezogen werden. Zur Präferenzmessung im Gesundheitswesen werden zunehmend Discrete-Choice-Experimente eingesetzt. Ein Vorteil dieser Methode ist die Fundierung in der mikroökonomischen Theorie. Großbritannien, Australien und Deutschland diskutieren derzeit Methoden und Verfahren für die Berücksichtigung der Patientenpräferenzen in der regulatorischen (Zulassung, Erstattung und Preisgestaltung) und gesundheitspolitischen Entscheidungsfindung. Allerdings bleibt offen, wie die Präferenzen der Patienten und der Öffentlichkeit am besten mit Expertenurteilen in Verbindung gebracht werden können.

Schlüsselwörter

Discrete-Choice-Experimente, Patientenpräferenzen, Nutzenbewertung

>> Der Begriff „Präferenz“ bezeichnet allgemein den Vorzug, Vorrang oder die Begünstigung einer Alternative bzw. die Vorliebe, die jemand für etwas hat. Als ein eindimensionaler Indikator zur Erklärung von (Wahl-) Entscheidungen repräsentieren Präferenzen das Ausmaß der Vorziehenswürdigkeit eines Beurteilungsobjekts für eine bestimmte Person während eines bestimmten Zeitraumes (Voeth 2000). Präferenzen sind damit die Grundlage für die Wahl einer Option; der Nutzen einer Option ist die Grundlage der Präferenz. Bekannte Präferenzen erlauben daher Rückschlüsse auf den individuellen Nutzen der zur Verfügung stehenden Optionen (Telser 2002).

Eine über die Messung der Patientenzufriedenheit hinausgehende Bewertung von Versorgungsprozessen müsste die Prioritäten der Patienten und die Vorziehenswürdigkeit der Alternativen zur Grundlage der Bewertung machen. Nur auf Basis der konkreten Wahlentscheidungen von Betroffenen können die Prioritäten auch gemessen werden. Diesbezügliche Wahlentscheidungen können sich auf die folgenden Sachverhalte beziehen:

- Patientenrelevante Endpunkte (Cost-Benefit Trade-off),
- Risiken/Sicherheit bei der Anwendung (Cost-Risk/Safety Trade-off) sowie
- Eigenschaften der Therapie (Properties Trade-off).

Einen methodischen Ansatz zur Entwicklung eines adäquaten Verfahrens bildet die im Rahmen der Präferenzforschung verwendete Discrete-Choice-Analyse, die auf Grundlage individueller Auswahlentscheidungen die Bewertung und Prognose von Präferenzen ermöglicht. Multiattributive Bewertungsverfahren, wie das Discrete-Choice-Experiment (DCE), beruhen auf der Auswertung von offenbarten Präferenzen (stated preferences). Es gibt zwei Arten von multiattributiven Bewertungsverfahren:

(1) präferenzbasierte Verfahren (preference-based), bei denen die Individuen aufgefordert werden, jede Produktalternative zu bewerten und/oder zu reihen.
(2) wahlbasierte Verfahren (choice-based), bei denen dem Befragten einige Produktalternativen vorgeben werden, aus denen dieser dann eine Alternative wählt.

Allgemein werden präferenzbasierte Verfahren mit dem Oberbegriff Conjoint-Analysen bezeichnet, während die Unterform der wahlbasierten Verfahren als Discrete-Choice-Analysen benannt wird. Das Discrete-Choice-Experiment ist eine häufig eingesetzte Methode für die Quantifizierung von Präferenzen der Bürger, Versicherten oder Patienten (Ryan et al. 2001; Marshall et al. 2010; de Bekker Grob et al. 2010). Discrete-Choice-Analysen sind dekompositionelle Methoden (Ryan et al. 2008), die durch die theoretischen Arbeiten von Lancaster (1966) und McFadden (1973) ermöglicht wurden. Erst diese theoretischen Arbeiten erlauben es, die Reihung (ranking) bzw. Bewertung (rating) von vielen verschiedenen Produkt- bzw. Dienstleistungsvarianten durch einen paarweisen Vergleich, d.h. durch eine (Aus-) Wahlentscheidung, zu ersetzen. Diese Methode der Präferenzmessung verwendet eine stochastische Nutzenfunktion (Random Utility Function), bei der die Zufallskomponente alle undefinierten Faktoren umfasst, welche die Wahl beeinflussen. Die Zufallsnutzentheorie (Random Utility Theory) geht davon aus, dass der Nutzen eines Gutes oder einer Dienstleistung aus einer systematischen oder beobachtbaren Komponente und einer zufälligen, nicht beobachtbaren Komponente besteht (Thurstone 1927; McFadden 1973; Manski 1977). DCEs beruhen auf der von Lancaster formulierten Nachfragetheorie, wonach der Nutzen, den ein Produkt oder eine Dienstleistung beim Konsumenten (bzw. Nachfrager oder Patienten) stiftet, von den Eigenschaften des Produktes und den Merkmalen des Konsumierenden abhängt. Der Nutzen einer Therapie ist demnach abhängig von den Eigenschaften der Therapie, z.B. Wirkung, Reduktion der Symptomlast oder Nebenwirkung etc.

Discrete-Choice-Experimente sind speziell entwickelt worden, um Informationen über die Bereitschaft eines Individuums, Kompromisse zwischen den einzelnen Eigenschaften (Attributen) eines Produktes zu akzeptieren, zu liefern. Es basiert auf dem hedonistischen Prinzip, dass Produkte aus einer Reihe von verschiedenen Attributen bestehen und dass die Attraktivität eines Produkts für ein Individuum eine Funktion aus diesen Attribute darstellt. DCEs beruhen auf der Annahme, dass Individuen sich in Wahlsituationen jeweils für die Alternative ent-

scheiden, die ihnen den höchsten Nutzen garantiert. Der Beobachter kennt diesen Nutzen aber nicht mit Sicherheit und muss sie deshalb als Zufallsvariable behandeln (Ben-Akiva/Lerman 1985; Anderson et al. 1992). Durch die statistische Analyse der individuellen Wahlentscheidungen kann die relative Bedeutung der Attribute und der Tradeoffs identifiziert werden (d.h. die Höhe einer Eigenschaft, die Befragte bereit sind aufzugeben, um mehr von einer anderen Eigenschaft zu bekommen). Die relativen Präferenzen der Individuen zwischen Produkteigenschaften und Levels variieren, und damit sind diese bereit, Kompromisse zu akzeptieren (Thurstone 1974; Johnson et al. 2000; Telser/Zweifel 2002; Hensher et al. 2005).

IQWiG-Pilotstudie: Das deutsche Methodenkonzept für die Kosten-Nutzen-Bewertung sieht die Verwendung einer Effizienzgrenze innerhalb eines Indikationsgebietes vor. Der Nutzen wird vom IQWiG anhand patientenrelevanter Endpunkte bewertet. Bei der Entscheidungsfindung auf Basis mehrerer Effizienzgrenzen bleibt jedoch offen, wie die Ergebnisse durch den Entscheidungsträger interpretiert werden. Das Ziel des IQWiG-Pilotprojekts bestand in der Darstellung der Nutzungsmöglichkeiten der Conjoint/Discrete-Choice-Analyse zur Identifikation patientenrelevanter Endpunkte. Die erste Fragestellung war, wie wir diese Entscheidungskriterien (Endpunkte) identifizieren werden können? Weiterhin sollte geprüft werden, ob diese Methode für die Gewichtung und Priorisierung über multiple Endpunkte anwendbar ist. Darauf aufbauend wurde geprüft, inwiefern diese Methode zur Priorisierung von Gesundheitstechnologien auf Basis multipler Endpunkte eingesetzt werden kann. Ein indikationsspezifischer Vergleich von Therapiealternativen ist denkbar. Zielsetzung der Machbarkeitsstudie waren die Identifikation, Gewichtung und Aggregation von patientenrelevanten Endpunkten am Beispiel der antiviralen Hepatitis C-Therapie (Mühlbacher et al. 2012).

Im Ergebnis zeigte sich, dass die Discrete-Choice-Methode eine theoretisch fundierte Methode zur Identifikation, Gewichtung und Priorisierung von patientenrelevanten Endpunkten bietet. Weiterhin können mit Hilfe von DCEs auch Präferenzunterschiede zwischen Experten und Patienten aufgezeigt und verglichen werden. Ferner kann die Wertfunktion approximiert und ein übergeordneter Nutzenwert gebildet werden. Die grafische Ableitung einer übergeordneten Effizienzgrenze ist somit darstellbar (Mühlbacher et al. 2011).

Diskussion: DCEs zur Identifizierung und Gewichtung Patientenrelevanter Endpunkte

Die Methode der Discrete-Choice-Experimente bietet zahlreiche praktische Vorteile (Ryan et al. 2008), so dass sie eine international etablierte Methode zur Erhebung von Präferenzen darstellt. Sie gilt auf der Grundlage jahrelanger Forschung als gut etabliert und findet in zunehmendem Maße im Gesundheitswesen Anwendung (Louviere et al. 2000; Ryan/Farrar 2000; Phillips et al. 2002). Die Anwendung der Conjoint-Analyse steigt rapide auch in der Medizin (Bridges et al. 2008), und Fachgesellschaften haben Leitlinien für die Anwendung der Conjoint-Analyse in der Versorgungsforschung entwickelt (Bridges et al. 2011).

Discrete-Choice-Experimente bieten umfassendste Informationen über die Austauschbeziehungen, die Individuen bereit sind, zwischen Wirksamkeit einer Behandlung, Sicherheit und anderen Faktoren, die das Wohlbefinden des Patienten bestimmen, einzugehen (Louviere et al. 2000; Flynn et al. 2008; Louviere/Lancsar 2009).

Ungeachtet der wachsenden Popularität von Stated-Preference-Methoden bei Gesundheitsökonomen sind diese noch nicht in einem ausreichenden Maß bei Entscheidungsträgern in der Industrie und bei Leistungserbringern bekannt. Sie wurden zudem bislang nur begrenzt beim Health Technology Assessment oder in der gesundheitsökonomischen Evaluation eingesetzt (Bridges 2003).

Allerdings werden die Anwendungsmöglichkeiten von Discrete-Choice-Modellen zur Ermittlung der Patientenpräferenzen und der Zahlungsbereitschaft im Rahmen der Kosten-Nutzen-Analyse und Kosten-Wirksamkeits-Analyse von medizinischen Interventionen zunehmend untersucht (Johnson/Hauber 2008; Lancsar/Louviere 2008; Mühlbacher et al. 2009; Mühlbacher et al. 2011; Mühlbacher et al. 2011).

Hintergrund: Patientenorientierung und patientenrelevante Endpunkte

Internationale Entwicklungen: In den letzten Jahren ist die Berücksichtigung der öffentlichen und Patienten-Präferenzen zunehmend diskutiert worden. Dies resultierte in einer Agenda für die Beteiligung der Öffentlichkeit in der Gesundheitsversorgung, nicht nur in Australien (Australian Research Council 2010), Kanada und Großbritannien, sondern auch in Deutschland. Die jüngsten Reformen versprechen ein positives Klima für eine Beteiligung der Öffentlichkeit in der Diskussion über vergleichende Effizienz, Sicherheit und Wirtschaftlichkeit / Wirksamkeit von Gesundheitsleistungen (Stafinski et al. 2011). In einigen Ländern werden die Meinungen der Öffentlichkeit sowie Patienten routinemäßig im Prozess des Health Technology Assessments und der Priorisierung berücksichtigt (National Institute for Health and Clinical Excellence 2012).

Der Patient als Experte: Es ist der Patient, der seine Präferenzen in Bezug auf bestimmte Behandlungen sowie Behandlungsergebnisse am besten kennt (Billings/Krakauer 2011). Schließlich ist es der Patient, der mit den Folgen von Behandlungsentscheidungen leben muss (Telser et al. 2008). Daher ist es wichtig, dass Entscheidungsträger sich der Präferenzen der Patienten bewusst sind und diese respektieren (Bleichrodt/Johannesson 1997). Patienten- oder öffentliche Präferenzen sollten auch auf der Makro-Ebene betrachtet werden und Entscheidungen über die Gestaltung der Gesundheitssysteme und die Allokation der begrenzten Ressourcen unterstützen (Stein 2004).

Entscheidungsfindung im Gesundheitswesen: Komplexe Entscheidungsprozesse im Gesundheitswesen erfordert oft die explizite Berücksichtigung mehrerer Entscheidungskriterien. Im Allgemeinen erfolgt die Entscheidungsfindung im Gesundheitswesen auf drei verschiedenen Ebenen und wirkt sich daher auf drei verschiedene Gruppen von Beteiligten aus.

- **Mikro-Ebene (partizipative Entscheidungsfindung):** Die Mikro-Ebene beinhaltet die Behandlungsentscheidungen für den einzelnen Patienten. Entscheidungen auf dieser Ebene werden auf Basis der Relevanz der Therapie für die Krankheit und der Auswirkungen auf die Lebensqualität des Patienten getroffen (Montgomery/Fahey 2001). Die Entscheidungen haben direkten Einfluss auf die Therapie (Eddy 2005).
- **Meso-Ebene (evidenzbasierte Medizin/EbM):** Die Meso-Ebene beschreibt die Entscheidungen für eine bestimmte Patientengruppe. Der Schwerpunkt der Entscheidungsfindung auf dieser Ebene ist die Definition der evidenzbasierten Medizin / Pflege durch evidenzbasierte (klinische) Richtlinien. Hier ist es wichtig, die Patienten- oder die öffentlichen Präferenzen sowie Expertenurteile (Ärzte) zu berücksichtigen (Haynes et al. 2002; van Til 2009).
- **Makro-Ebene (öffentliche Präferenzen und regulatorische Entscheidungen):** In der evidenzbasierten Gesundheitsversorgung werden die Prinzipien der EbM an die organisatorischen und institutionellen Ebenen kommuniziert. Dies bedeutet, dass die Therapieentscheidungen nicht für einen einzelnen Patienten oder einer Gruppe von Patienten, sondern für eine ganze Bevölkerung getroffen werden (Eddy 2005). Die wichtigsten Ansätze auf dieser Ebene sind Risiko-Nutzen-oder Kosten-Nutzen-Analysen.

Patientenrelevante Endpunkte: Patientenrelevante Endpunkte können Mortalität, Morbidität und gesundheitsbezogener Lebensqualität sowie andere wichtige Faktoren beinhalten, die in der Analyse als patientenrelevant identifiziert wurden. Es ist wichtig festzustellen, welche Faktoren die relevantesten für die Patienten sind. Eine einfache gleiche Gewichtung von Endpunkten bei der Schaffung einer Aggregation ist suboptimal. Der Nutzen, im Sinne eines aggregierten (Gesamt-) Patientennutzens, kann mithilfe unterschiedlicher patientenrelevanter Endpunkte nur sehr schwer abgebildet werden.

Trotz der spezifischen Besonderheiten der Gesundheitsmärkte ist eine Vielzahl an Studien zu den Präferenzen im Hinblick auf die Leistungserbringung, Therapieeigenschaften und der Compliance durchgeführt worden.

Es existieren heute unter anderem Untersuchungen, die auf dieser Vorgehensweise beruhen, zu:

- Psoriasis (Ashcroft et al. 2006; Seston et al. 2007),
- Übergewicht (Roux et al. 2004),
- Chronischen Schmerzen (Mühlbacher et al. 2011),
- Multiplem Myelom (Mühlbacher et al. 2008; Mühlbacher/Nübling 2011),
- HIV/AIDS (Mühlbacher et al. 2011; Mühlbacher et al. 2012),
- ADHS (Mühlbacher et al. 2009; Mühlbacher/Nübling 2010),
- Multipler Sklerose (Johnson et al. 2009),
- der Lebensqualität bei Krebspatienten (Johnson et al. 2006),
- Morbus Crohn (Johnson et al. 2010)
- Hepatitis B (Mohamed et al. 2012)
- Nicht-kleinzelligem Lungenkarzinom (Bridges et al. 2012) oder
- Typ II-Diabetes (Hauber et al. 2009),

Grundsätzlich können Patientenpräferenzen auf allen Ebenen in die Entscheidungsprozesse einbezogen werden. Im internationalen Kontext wird zunehmend diskutiert, wie die Perspektive der betroffenen Konsumenten berücksichtigt werden kann. Die Spannbreite der Partizipationsbemühungen erstreckt sich über die qualitative Erhebung von Patientenbedürfnissen (Citizen Councils des NICE in Großbritannien) bis hin zur wissenschaftlich fundierten Dokumentation quantitativer Patientenpräferenzen (Patient-Centered Outcomes Research des PCORI in den USA). Kanada und Australien arbeiten auch an Reformen, welche die Einbeziehung der Patienten in die Entscheidungsprozesse begünstigen sollen. International werden zunehmend die Methoden der multikriteriellen oder multiattributiven Entscheidungsfindung (Multi-Criteria Decision Making) als Alternativen zum Konzept der Kosten-Effektivitäts-Studien (Cost-Effectiveness Studies) diskutiert. Allen Ansätzen ist gemein, dass der Patientennutzen (Gesamtnutzen) als multidimensionales Konstrukt angesehen wird, welcher nur unter Einbeziehung der betroffenen Patienten bewertet werden kann. Als eindimensionaler Indikator zur Erklärung von (Wahl-) Entscheidungen repräsentieren Patientenpräferenzen das Ausmaß der Vorziehenswürdigkeit einer Eigenschaft oder eines Bewertungsgegenstandes. Der wahrgenommene oder erwartete Nutzen aus Sicht des Patienten ist damit die Basis für Präferenzen.

Literatur (Teil 1)

Anderson, S. P., De Palma, A., Thisse, J. F. (1992): Discrete choice theory of product differentiation, MIT press.

Ashcroft, D. M., Seston, E., Griffiths, C. E. (2006): „Trade-offs between the benefits and risks of drug treatment for psoriasis: a discrete choice experiment with U.K. dermatologists.“ Br J Dermatol 155(6): 1236-1241.

Australian Research Council (2010): Summary of Successful Linkage - Projects Proposals for Funding to Commence in 2010 by State and Organisation.

Ben-Akiva, M., Lerman, S. (1985): Discrete choice analysis: theory and application to travel demand, MIT press.

Billings, J. A., Krakauer, E. L. (2011): „On patient autonomy and physician responsibility in end-of-life care.“ Arch Intern Med 171(9): 849-853.

Bleichrodt, H., Johannesson, M. (1997): „Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs.“ Journal of Health Economics 16(2): 155-175.

Bridges, J. (2003): „Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics.“ Applied Health Economics and Health Policy 2(4): 213-224.

Bridges, J., Hauber, B., Marshall, D., Lloyd, A., Prosser, L. A., Regier, D. A., Johnson, R. F., Mauskopf, J. (2011): „Conjoint Analysis Use in Health Studies – A Checklist: A Report of the ISPOR Conjoint Analysis in Health Good Research Practices Task Force.“ ISPOR TF Report.

Bridges, J. F., Kinter, E. T., Kidane, L., Heinzen, R. R., McCormick, C. (2008): „Things are Looking up Since We Started Listening to Patients: Trends in the Application of Conjoint Analysis in Health 1982-2007.“ The Patient: Patient-Centered Outcomes Research 14(4): 273-282.

Bridges, J. F. P., Mohamed, A. F., Finnern, H. W., Woehl, A., Hauber, A. B. (2012): „Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: A conjoint analysis.“ Lung Cancer.

de Bekker Grob, E. W., Ryan, M., Gerard, K. (2010): „Discrete choice experiments in health economics: a review of the literature.“ Health economics.

Eddy, D. M. (2005): „Evidence-based medicine: a unified approach.“ Health affairs 24(1): 9-17.

Flynn, T. N., Louviere, J. J., Peters, T. J., Coast, J. (2008): „Estimating preferences for a dermatology consultation using Best-Worst Scaling: comparison of various methods of analysis.“ BMC Med Res Methodol 8: 76.

Green, C., Gerard, K. (2009): „Exploring the social value of health-care interventions: a stated preference discrete choice experiment.“ Health economics 18(8): 951-976.

Hart, D. (2005): „Die Nutzen/Risiko-Abwägung im Arzneimittelrecht.“ Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz 48(2): 204-214.

Hauber, A., Mohamed, A., Johnson, F., Falvey, H. (2009): „Treatment preferences and medication adherence of people with Type 2 diabetes using oral glucose-lowering agents.“ Diabetic Medicine 26(4): 416-424.

Haynes, R. B., Devereaux, P., Guyatt, G. H. (2002): „Physicians' and patients' choices in evidence based practice.“ BMJ 324(7350): 1350.

Hensher, D., Rose, J., Greene, W. (2005): Applied choice analysis: a primer, Cambridge Univ Pr.

Johnson, F., Hauber, A., Osoba, D., et al. (2006): „Are chemotherapy patients' HRQoL importance weights consistent with linear scoring rules? A stated-choice approach.“ Quality of Life Research 15(2): 285-298.

Johnson, F. R., Banzhaf, M. R., Desvousges, W. H. (2000): „Willingness to pay for improved respiratory and cardiovascular health: a multiple-format, stated-preference approach.“ Health Econ 9(4): 295-317.

Johnson, F. R., Hauber, B., Ozdemir, S., Siegel, C. A., Hass, S., Sands, B. E. (2010): „Are gastroenterologists less tolerant of treatment risks than patients? Benefit-risk preferences in Crohn's disease management.“ J Manag Care Pharm 16(8): 616-628.

Johnson, F. R., Van Houtven, G., Özdemir, S., Hass, S., White, J., Francis, G., Miller, D. W., Phillips, J. T. (2009): „Multiple sclerosis patients—benefit-risk preferences: Serious adverse event risks versus treatment efficacy.“ Journal of neurology 256(4): 554-562.

Johnson, R., Hauber, B. (2008): Quantifying Patient Benefit-Risk Tradeoff Prefe- rences: A Brief Introduction, RTI Health Solutions- White Paper.

Klinke, A., Renn, O. (2002): „A New Approach to Risk Evaluation and Management: Risk-Based, Precaution Based, and Discourse Based Strategies1.“ Risk analysis 22(6): 1071-1094.

Lancaster, K. J. (1966): „A new approach to consumer theory.“ The journal of political economy 74(2): 132-157.

Lancsar, E., Louviere, J. (2008): „Conducting Discrete Choice Experiments to Inform Healthcare Decision Making: A Users Guide.“ Pharmacoeconomics 26(8): 661-677.

Louviere, J., Hensher, D., Swait, J. (2000): Stated choice methods: analysis and applications, Cambridge Univ Press.

Louviere, J. J., Lancsar, E. (2009): „Choice experiments in health: the good, the bad, the ugly and toward a brighter future.“ Health Econ Policy Law 4(Pt 4): 527-546.

Manski, C. F. (1977): „The structure of random utility models.“ Theory and decision 8(3): 229-254.

Marshall, D., Bridges, J. F. P., Hauber, B., Cameron, R., Donnalley, L., Fyie, K., Johnson, F. R. (2010): „Conjoint Analysis Applications in Health How are Studies being Designed and Reported?: An Update on Current Practice in the Published Literature between 2005 and 2008.“ The Patient: Patient-Centered Outcomes Research 3(4): 249-256.

McFadden, D. (1973): „Conditional logit analysis of qualitative choice behavior.“

McFadden, D. (1974): „Conditional logit analysis of qualitative choice behavior.“ Zarembka 1974: 105-142.

Vor dem Hintergrund der Einbeziehung von Patientenpräferenzen im Rahmen der Zulassung, Beurteilung der Erstattungsfähigkeit und Preisfestsetzung von Arzneimitteln stellt sich für die Öffentlichkeit die Frage „Wie unsicher ist sicher genug?“, während für den individuellen Patienten die Frage „Welche Risiken nehme ich für welchen Nutzen auf mich?“ von großer Bedeutung ist (Klinke/Renn 2002). Im Rahmen des klinischen und ökonomischen Entscheidungskalküls ist die Frage, wie die Interessen der Patienten in die Nutzen-Risiko-Abwägung integriert werden können. Hierbei ist zu beachten, dass die Perspektiven und Risikoeinstellungen von politischen Entscheidern, Medizinern und Patienten unterschiedlich sein können (Hart 2005).

Im Allgemeinen scheinen drei Einsatzbereiche für die Ergebnisse aus Präferenzmessungen denkbar:

(1) Utility Assessment/Health Technology Assessment: Beschreibung des Nutzens unterschiedlicher Alternativen durch die Addition der vorliegenden Präferenzgewichte einer real existierenden Alternative.

(2) Produktdesign & Evaluation: Informationen darüber, wie bestimmte Produkte ausgestaltet sein müssten, damit sie für bestimmte Patientengruppen den höchsten Nutzen erzeugen.

(3) Prognose: Bereitstellung von prognostizierenden Informationen darüber, wie stark ein hypothetisches Produkt die Akzeptanz der Patienten beeinflussen würde.

Die Methode eignet sich darüber hinaus zur Bewertung von Gesundheitszuständen (health valuation) sowie zur Operationalisierung eines multidimensionalen Gesamtnutzens bzw. der Erhebung der Zahlungsbe-

Literatur (Zeil 2)

Mohamed, A. F., Johnson, F. R., Hauber, A. B., Lescrauwaet, B., Masterson, A. (2012): „Physicians' stated trade-off preferences for chronic hepatitis B treatment outcomes in Germany, France, Spain, Turkey, and Italy.“ European journal of gastroenterology & hepatology 24(4): 419.

Montgomery, A., Fahey, T. (2001): „How do patients' treatment preferences compare with those of clinicians?“ Quality in Health Care 10(suppl 1): i39-i43.

Mühlbacher, A. (2012): Alle Mühen umsonst? Studie: Patienten haben wenig Interesse an Vernetzung. 7. Kongress für Gesundheitsnetzwerker 2012, Berlin, Kongressinitiative der Berlin-Chemie AG.

Mühlbacher, A., Bethge, S., Schulman, K. (2011): „Patient-Centered Health Care Delivery Systems: A Discrete-Choice Experiment.“ Value in Health 14(7): A349.

Mühlbacher, A., Bethge, S., Tockhorn, A. (2012): „Präferenzmessung im Gesundheitswesen: Grundlagen von Discrete-Choice-Experimenten.“ Gesundh ökon Qual manag 17: 1-15 (in press).

Mühlbacher, A., Bridges, J., Bethge S, et al. (2012): Pilotstudie Wahlbasierte Conjoint Ananlyse in der Kosten-Nutzen-Bewertung: Ermittlung eines endpunktbasierten Nutzenwertes. unveröffentlichter Endbericht, Hochschule Neubrandenburg, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Köln.

Mühlbacher, A., Bridges, J. F. P., Bethge, S. (2011): „Prioritization and weighting of patient-relevant endpoints as part of the IQWiGs efficiency frontier method in Germany.“ Value in Health 14(3): A73 - A74.

Mühlbacher, A., Ezernieks, J., Nübling, M. (2011): „Chronic Pain: Patient Treatment Preferences - A Discrete Choice Experiment.“ Value in Health 14(7): A417.

Mühlbacher, A., Juhnke, C. (2012): „Patients Preferences versus Physicians Judgments: Does it make a difference in health care decision-making?“ Applied Health Economics and Health Policy(in press).

Mühlbacher, A., Lincke, H. J., Nübling, M. (2008): „Evaluating patients' preferences for multiple myeloma therapy, a Discrete-Choice-Experiment.“ GMS Psycho-Social Medicine 5(Doc10): 1-10.

Mühlbacher, A., Nübling, M. (2010): „Analysis of Patients Preferences: Direct Assessment and Discrete-Choice Experiment in Therapy of Adults with Attention-Deficit Hyperactivity Disorder.“ The Patient: Patient-Centered Outcomes Research 3(4): 285-294.

Mühlbacher, A., Nübling, M. (2011): „Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma.“ The European Journal of Health Economics 12(3): 193-203.

Mühlbacher, A., Rudolph, I., Lincke, H. J., Nübling, M. (2009): „Preferences for treatment of Attention-Deficit/Hyperactivity Disorder (ADHD): a discrete choice experiment.“ BMC health services research 9(1): 149.

Mühlbacher, A., Stoll, M., Mahlich, J., Nübling, M. (2012): Evaluating Patients' Preferences for AIDS/HIV Therapy - A Discrete-Choice-Experiment. HIV medicine (in review).

Mühlbacher, A., Stoll, M., Ranneberg, B., Nübling, M. (2011): „Discrete Choice Experiments (DCEs) In HIV/Aids Treatment: Expert Judgment In Comparison To Patient Preferences.“ Value in Health 14(7): A282.

Mühlbacher, A. C., Nübling, M. (2011): „Analysis of physicians' perspectives versus patients' preferences: direct assessment and discrete choice experiments in the therapy of multiple myeloma.“ The European Journal of Health Economics 12(3): 193-203.

National Institute for Health and Clinical Excellence. „Patient and public involvement policy.“ Retrieved http://www.nice.org.uk/getinvolved/patientandpublicinvolvement/patient_and_public_involvement.jsp., 2012.

Phillips, K. A., Johnson, F. R., Maddala, T. (2002): „Measuring what people value: a comparison of „attitude“ and „preference“ surveys.“ Health Serv Res 37(6): 1659-1679.

Ratcliffe, J. (2000): „Public preferences for the allocation of donor liver grafts for transplantation.“ Health economics 9(2): 137-148.

Roux, L., Ubach, C., Donaldson, C., Ryan, M. (2004): „Valuing the benefits of weight loss programs: an application of the discrete choice experiment.“ Obesity research 12(8): 1342-1351.

Ryan, M., Bate, A., Eastmond, C., Ludbrook, A. (2001): „Use of discrete choice experiments to elicit preferences.“ Quality in Health Care 10(suppl 1): i55-i60.

Ryan, M., Farrar, S. (2000): „Using conjoint analysis to elicit preferences for health care.“ Bmj 320(7248): 1530-1533. Ryan, M., Bate, A., Eastmond, C., Ludbrook, A. (2001): „Use of discrete choice experiments to elicit preferences.“ Quality in Health Care 10(suppl 1): i55-i60.

Ryan, M., Farrar, S. (2000): „Using conjoint analysis to elicit preferences for health care.“ Bmj 320(7248): 1530-1533.

Ryan, M., Gerard, K., Amaya-Amaya, M. (2008): Using discrete choice experiments to value health and health care, Springer Academic Publishers.

Seston, E. M., Ashcroft, D. M., Griffiths, C. E. (2007): „Balancing the benefits and risks of drug treatment: a stated-preference, discrete choice experiment with patients with psoriasis.“ Arch Dermatol 143(9): 1175-1179.

Stafinski, T., D. Menon, D.J. Philippon, McCabe, C. (2011): „Health Technology Funding Decision -Making Processes Around the World: The Same, Yet Different.“ PharmacoEconomics 29(6): 475-495.

Stein, J. D. (2004): „Disparities between ophthalmologists and their patients in estimating quality of life.“ Current opinion in ophthalmology 15(3): 238.

Telser, H. (2002): Nutzenmessung im Gesundheitswesen: Die Methode der Discrete-Choice-Experimente, Verlag Dr. Kova-.

Telser, H., Becker, K., Zweifel, P. (2008): „Validity and reliability of willingness-to-pay estimates: evidence from two overlapping discrete-choice experiments.“ The Patient: Patient-Centered Outcomes Research 1(4): 283-298.

Telser, H., Zweifel, P. (2002): „Measuring willingness-to-pay for risk reduction: an application of conjoint analysis.“ Health Econ 11(2): 129-139.

Thurstone, L. L. (1927): „A law of comparative judgment.“ Psychological Review; Psychological Review 34(4): 273.

Thurstone, L. L. (1974): „A law of comparative judgment.“ Scaling: A sourcebook for behavioral scientists: 81-92.

van Til, J. A. (2009): Integrating preferences into decision making - The treatment of ankle-foot impairment in stroke, Dissertation.

Voeth, M. (2000): Nutzenmessung in der Kaufverhaltensforschung: Die hierarchische individualisierte Limit-Conjoint-Analyse (HILCA), Deutscher Universitäts-Verlag.

Whitty, J. A., Scuffham, P. A., Rundle-Thiele, S. R. (2011): „Public and decision maker stated preferences for pharmaceutical subsidy decisions: a pilot study.“ Applied health economics and health policy 9(2): 73-79.

reitschaft als zusätzliches Entscheidungskriterium im Rahmen des Health Technology Assessment (HTA), so z. B. in den unterschiedlichen Phasen der Entwicklung von Arzneimittelinnovationen bzw. Gesundheitstechnologien (Mühlbacher et al. 2012). Die Anwendung von Präferenzdaten im Produktdesign wurde im Rahmen der Beschreibung eines optimalen Versorgungsprogramms bereits getestet. In zwei ähnlichen Studien in den USA (Mühlbacher et al. 2011) und in Deutschland (Mühlbacher 2012) wurden beispielsweise 21 Attribute von integrierten Versorgungsnetzwerken patientenspezifisch bewertet. Bei einer Studie, die sich dem dritten Einsatzgebiet widmet, läge der Fokus hier auf der Compliance oder Adhärenz: Wie kann durch präferenzbasierte Versorgungsangebote das Verhalten der Patienten gesteuert werden? Neben der Erfassung von Patientenpräferenzen wurde das Discrete-Choice-Experiment bereits angewandt, um soziale Präferenzen zu erheben (Ratcliffe 2000; Green/Gerard 2009; Whitty et al. 2011). Soziale Präferenzen werden unter der impliziten Annahme gemessen, dass Individuen bei ihrer Entscheidung altruistisch handeln, um ihren individuellen Nutzen durch die Maximierung des gesellschaftlichen Nutzens als Ganzes zu maximieren. Weiterhin können mit Hilfe von Discrete-Choice-Experimenten auch Präferenzunterschiede zwischen Experten und Patienten aufgezeigt und verglichen werden (Mühlbacher/Nübling 2011; Mühlbacher/Juhnke 2012).

Zusammenfassung

Mit der Einbeziehung bzw. Berücksichtigung von klinischen Endpunkten in die ökonomische Nutzenbewertung werden zwei derzeit unabhängige Studienwelten vereint. Zukünftig könnten Studienergebnisse auf Basis der evidenzbasierten Medizin analysiert und im Kontext der Medizin und Ökonomie interpretiert werden. Der Staat, Behörden oder öffentlich-rechtliche Institutionen als Sachwalter der Bürger, Versicherten oder Patienten entscheiden über die Ausgestaltung des Versorgungssystems und des Leistungskataloges. Hierzu bedarf es der Bewertung und Evaluation von Leistungsangeboten und der Prognose von Nachfrage und Akzeptanz der zur Verfügung stehenden Alternativen. Die Erstattung bzw. Erstattungsfähigkeit von Leistungen und Produkten ist ein zentraler Bereich, über den entschieden werden muss. Bei der Kostenübernahme gilt es zu bestimmen, ob es sinnvoll ist, einen höheren Preis für ein neues Produkt zu akzeptieren. Ein höherer Preis ist nur dann gerechtfertigt, wenn mit einer Leistung auch ein höherer Patientennutzen bzw. ein geringerer Schaden für den Patienten verbunden ist. Wenn ein Anbieter ein Verfahren, ein Produkt oder eine Technologie im deutschen Gesundheitssystem einbringen will, geht es primär um die Frage „Was hat der Patient davon?“ Damit wird deutlich, dass die Präferenzen als Ausdruck des Patientennutzens eine entscheidende Rolle bei der Entscheidungsfindung im Gesundheitswesen spielen müssen. Eine Methode zur Erfassung und Dokumentation von latenten Präferenzstrukturen und den Abwägungsprozessen in Bezug auf Behandlungsmethoden ist die ökonomisch fundierte Discrete-Choice-Analyse (McFadden 1974; Thurstone 1974), mit deren Hilfe der patientenseitige Zusatznutzen von Produkten, Dienstleistungen oder innovativen Technologien nachgewiesen werden kann. <<

Health Technology Assessment from the patient perspective?

The documented patients benefit should be decisive in priority setting. The underlying decision-making processes must ensure that the selection of a variety of potential health goods and services is in line with the needs, expectations and priorities of the insured and patients. The common basis of the clinical and economic decision-making calculus is the benefit to patients. Therefore patient preferences can and should be involved at all levels in the decision-making processes. Great Britain, Australia and Germany are currently discussing methods and procedures for valuing patient preferences in regulatory (authorization and pricing) decision-making. However, many questions remain on how to best balance patient and public preferences with physicians' judgment in healthcare and health policy decision-making. For example, how to define evaluation criteria, regarding the perceived value from a patient's perspective?

Keywords

Discrete-Choice Experiments, Patient Preferences, Utility Assessment

Prof. Dr. Axel C. Mühlbacher
ist Professor an der Hochschule Neubrandenburg, IGM Institut Gesundheitsökonomie und Medizinmanagement sowie Mitgründer der Gesellschaft für empirische Beratung mbH.

Kontakt: muehlbacher@hs-nb.de



Susanne Bethge, M.Sc.
ist wissenschaftliche Mitarbeiterin an der Hochschule Neubrandenburg im Institut Gesundheitsökonomie und Medizinmanagement (IGM) sowie Doktorandin an der Medizinischen Hochschule Hannover.

Kontakt: bethge@hs-nb.de.



Christin Juhnke, M.A.
ist wissenschaftliche Mitarbeiterin an der Hochschule Neubrandenburg im Institut Gesundheitsökonomie und Medizinmanagement (IGM).

Kontakt: juhnke@hs-nb.de



Dr. rer. biol. hum. Ch.-M. Dintsios, MPH

Patientengewichtete Endpunkte: Conjoint Analyse (CA) versus Analytic Hierarchy Process (AHP)

Die so genannte Multikriterielle Entscheidungsanalyse (MKEA) kann zur Unterstützung komplexer Technologiebewertungen eingesetzt werden. Ihren Ursprung haben multikriterielle entscheidungsanalytische Verfahren in der Entscheidungstheorie bzw. Operations Research und werden dort eingesetzt, um Entscheidungen unter Unsicherheit zu treffen. Sie unterstützen Entscheidungsträger bei der Bewertung einer endlichen Anzahl alternativer Gesundheitsinterventionen unter Verwendung einer endlichen Anzahl von Leistungskriterien. Zu den am meisten verbreiteten multikriteriellen Entscheidungsanalyseinstrumenten gehört neben der Conjoint-Analyse, dem Elimination and Choice translating Reality-Verfahren (ELECTRE) und der Multiattributiven Nutzwerttheorie (MAUT) (Lootsma/Schujit 1997) der Analytic Hierarchy Process (AHP).

>> Diese Instrumente unterscheiden sich hauptsächlich in ihren Ansätzen, Gewichte für Attribute der untersuchten Technologien zu erheben und präsentieren unterschiedliche statistische Modelle bei der Strukturierung der Präferenzen der Entscheidungsträger zu den zur Verfügung stehenden Alternativen. In dem von uns betrachteten Anwendungsbereich wären primär die befragten Patienten Entscheidungsträger und die jeweiligen Endpunkte innerhalb des betrachteten Krankheitsbildes die Leistungskriterien der zur Anwendung kommenden Gesundheitsinterventionen. Der Theorie nach liegt der Mehrwert dieser Instrumente in der Unterstützung der klinischen bzw. (gesundheits)ökonomischen Bewertung der Gesundheitstechnologien, die als multifaktorieller Prozess in entsprechende Entscheidungsfindungen eingehen.

Nichtsdestotrotz wird die Anwendbarkeit und Validität multikriterieller Entscheidungsanalyse hinterfragt. Beispielsweise liefern verschiedene multikriterielle Entscheidungsanalyseverfahren unterschiedliche Ergebnisse bei Anwendung im Rahmen desselben Entscheidungsproblems (Zanakis et al. 1998). Des Weiteren bleibt ihre prädiktive Validität oft genug spekulativ. Unternommene experimentelle Vergleiche schlussfolgerten, dass jede multikriterielle Entscheidungsanalysemethode ihre eigenen Vor- und Nachteile hat. Diese gilt es in vorliegender Abhandlung für die zwei etablierten Instrumente CA und AHP, die auch im Gesundheitsbereich (de Bekker-Grob et al. 2010; Wallenius et al. 2008) bzw. bei der Erhebung von Patientenpräferenzen immer mehr eingesetzt werden, kurz zu beleuchten und gegenüberzustellen.

Mathematisch-theoretische Grundlagen des AHP

AHP als kompositionelles Verfahren

Die Grundannahme des AHP lautet, dass der berechnete rechte Eigenvektor der zugrundegelegten Matrix A ungefähr dem Vektor der relativen Gewichte entspricht. Es handelt sich somit um ein kompositionelles Verfahren, das einem gewichteten additiven Nutzenmodell folgt (Abb. 1).

Zusammenfassung

Beide Verfahren, Conjoint Analyse (CA) und Analytic Hierarchy Process (AHP), werden zur Erhebung von Patientenpräferenzen (hier Priorisierung und Gewichtung von Endpunkten) eingesetzt. Über ihre theoretische Verortung hinaus (ökonomische Nutzentheorie versus deskriptive Theorie) unterscheiden sie sich hinsichtlich des zugrundegelegten Modells (additives Teilwertmodell versus gewichtetes additives Nutzenmodell), methodischen Ansatzes (dekompositionell versus kompositionell), Bewertungsprozesses (holistisch versus Paarvergleiche), methodischer Schwächen (kognitive Überforderung versus unrealen Entscheidungssituationen) und des Ressourcenaufwands (umfangreich versus gering) nicht unerheblich. CA kann bezogen auf a priori Annahmen zu widersprüchlichen Ergebnissen führen, AHP leidet unter dem so genannten Rank Reversal. Sie generieren allerdings in vergleichenden Studien durchaus ähnliche Ergebnisse und besitzen je nach Fragestellung ihre Vor- und Nachteile.

Schlüsselwörter

Multikriterielle Entscheidungsanalyse, Conjoint Analyse, Analytical Hierarchy Process, Patientenpräferenzen, Endpunktgewichtung, Nutzen-Schaden-Abwägung, Nutzenaggregierung

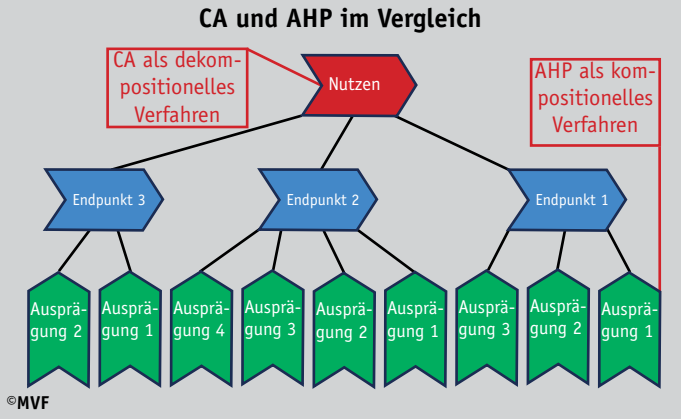


Abb. 1: Dekompositionelle CA versus Kompositioneller AHP.

Rechter Eigenvektor

Das AHP bietet einen Ansatz an, der zur Abschätzung von Gewichten der einzelnen Entscheidungskriterien und zur Priorisierung von zur Verfügung stehenden Alternativen geeignet ist. Diese Quantifizierung wird aus einer Matrix von paarweisen Vergleichen zweier Kriterien bzw. Alternativen hergeleitet. Der paarweise Vergleich der Entscheidungskriterien zeigt die relative Wichtigkeit dieser Kriterien bei der Zielerreichung, der Vergleich der Alternativen die relative Präferenz für die betrachteten Alternativen in Abhängigkeit von einem entsprechenden Entscheidungskriterium. Beide - relative Wichtigkeit der Kriterien und relative Präferenz der Alternativen - beruhen auf einer ordinalen Skala, die den numerischen Wertebereich von 1 für gleiche Wichtigkeit bzw. Präferenz bis zu 9 für extreme Wichtigkeit oder Präferenz innerhalb eines paarweisen Vergleichs einnimmt. Dabei stehen die einzelnen numerischen Ausprägungen der Ordinalskala für folgende Ausprägungen der relativen Wichtigkeit oder Präferenz:

- 1 für indifferente bzw. gleich wichtige Aussagen (Kriterien),
- 3 für etwas wichtiger,
- 5 für viel wichtiger,
- 7 für sehr viel wichtiger und
- 9 für extrem viel wichtiger.
- 2, 4, 6 und 8 werden jeweils als intermediäre Werte eingesetzt.

Die innerhalb des paarweisen Vergleichs betrachteten Kriterien können umgedreht werden, wenn die Richtung der Wichtigkeit sich umkehrt. Es entsteht eine Matrix von paarweisen Vergleichen über n Kriterien, mit A = (aij):

Matrix von paarweisen Vergleichen				
Kriterien	**Kriterium1**	**Kriterium 2**	**...**	**Kriterium n**
Kriterium 1	a_{11}	a_{12}	...	a_{1n}
Kriterium 2	a_{21}	a_{22}	...	a_{2n}
Kriterium 3	a_{31}	a_{32}	...	a_{3n}
...	...		...	
....	...		...	
Kriterium n	a_{n1}	a_{n2}	...	a_{nn}

Tab. 1: Matrix von paarweisen Vergleichen.

Sowohl Reflexivität (aij = 1 auf der Diagonalen) als auch Reziprozität (aij = 1/aji) werden in dieser Matrix vorausgesetzt. Aus diesem Grund müssen nur die Werte über die Diagonale eingetragen werden. Dies bedeutet, dass für n Kriterien nur n(n-1)/2 distinkte paarweise Vergleiche notwendig sind. Wenn im Rahmen des AHP Gruppen mit x Beteiligten befragt werden (x > 1), können die einzelnen Zellen der Matrix mit aggregierten paarweisen Vergleichen ausgefüllt werden, in dem das geometrische Mittel der individuellen Einzelbewertungen gebildet wird:

aij = (aij(1) * aij(2) * ... * aij(x))(1/x) für alle i,j = 1, 2, ..., n.

Wenn alle paarweisen Vergleiche in perfekter Konsistenz zu den verbleibenden Vergleichen stehen, dann gilt die Transitivität, formal ausgedrückt wie folgt:

aik * akj = aij für alle i, j, k = 1, 2, ..., n.

Das AHP liefert ein Konsistenzmaß, um auf mögliche Inkonsistenzen beim Kriterienvergleich hinzuweisen. Dieses Maß fußt auf dem maximalen Eigenwert λ_{max} der Matrix mit den paarweisen Vergleichen. Dieser Eigenwert ist immer größer oder gleich n für positive, reziproke Matrizen und gleich n, wenn die Matrix der paarweisen Vergleiche konsistent ist. Durch Normalisierung mit der Matrixgröße definiert Saaty den sogenannten Konsistenzindex C.I. in folgender Weise:

$$\mathbf{C.I. = (\lambda_{max} - 1)/(n - 1).}$$

Saaty (1989) generiert für eine Matrix der Größe n in einem Randomisierungsverfahren zufällig 500 Matrizen und berechnet den durchschnittlichen Konsistenzindex C.I. aus diesen Matrizen, den er Random Index (R.I.) nennt. Die Konsistenzratio (C.R.) wiederum gibt an, wie weit die paarweisen Bewertungen von einer völlig zufälligen Matrix paarweiser Vergleiche abweichen und wird definiert als Verhältnis zwischen Konsistenzindex und Random Index: C.R. = C.I./R.I.

Nach Saaty sind 10% der Inkonsistenz der zufälligen Matrix erlaubt. Dies bedeutet, dass ein Wert für die Konsistenzratio von unter 10% (C.R. < 0,1) toleriert bzw. akzeptiert werden kann. Derselbe Eigenwert der Matrix mit den paarweisen Vergleichen wird nun eingesetzt, um die Gewichte der Kriterien abzuschätzen:

$$\mathbf{wi = (\Sigma n\ j=1\ aij * wj) / \lambda_{max} \text{ für alle } I = 1, 2, ..., n.}$$

Diese Eigenvektormethode kann interpretiert werden als einfacher Mittelungsprozess, bei welchem die finalen Gewichte die Durchschnitte aller Vergleichsmöglichkeiten der Wichtigkeit der zu bewertenden Kriterien darstellen.

Die CA als dekompositionelles Verfahren

Die CA baut auf die Werttheorie nach Lancaster auf (Lancaster 1966). Der Nutzen eines „Gutes“ leitet sich aus den Eigenschaften und nicht aus dem „Gut“ selbst ab. Sie folgt somit der Theorie des stochastischen Nutzens (Random Utility Theory), wonach das Prinzip der Nutzenmaximierung gilt: Individuen entscheiden sich jeweils für die Alternative mit dem voraussichtlich höchsten Nutzen. Bei unvollständigen Informationen kann der Nutzen nicht vollständig erfasst werden. Er setzt sich additiv zusammen aus der beobachtbaren deterministischer Komponente (z.B. Merkmale/Attribute der Technologie, persönliche Eigenschaften) in Form einer linearen Nutzenfunktion (Schätzung der Gewichtung der Attribute) und aus einer nicht beobachtbaren stochastischer (Zufalls-)Komponente. Je nach Verteilungsannahmen erfolgt die Schätzung über ein Probit- oder Logit-Modell. Es handelt sich somit um ein dekompositionelles Verfahren, das einem additiven Teilwertmodell folgt (Tab. 1).

Logistische Regression

Die Conjoint Analyse folgt in der hier primär im Fokus stehenden Discrete Choice Variante einem (logistischen) Regressionsmodell, wonach für den zu ermittelnden Nutzen die Beziehung gilt:

$$\mathbf{U_{ij} = V_{ij} + \varepsilon_{ij}}$$

(U_{ij} = Nutzen („utility“), ε_{ij} = stochastische unbeobachtbare Komponente (u.a. Unterbewusstsein, körperliche Verfassung, Gemütszustand), V_{ij} = deterministische Komponente (Merkmale des Gutes/der Leistung und des Individuums), X = Merkmale (Attribute) des Gutes/der Leistung j; z = Merkmale des Individuums i; β = Maß für die Stärke der Gewichtung der Merkmale (Attribute) des Gutes/der Leistung j und γ = Maß für die Stärke der Gewichtung der Merkmale des Individuums i.

Die Gleichungen $\mathbf{V_{ij} = V\ (X_{ij,}\ z_{ij})}$ und $\mathbf{V_{ij} = \beta X_{ij} + \gamma z_{ij}}$ setzen dann die deterministische Komponente in Relation zu den Attributen des Gutes und den Merkmalen des Individuums.

Die Schätzung der relativen Bedeutung (Gewichtung) eines Attributes (hier Endpunktes) erfolgt dann im Rahmen einer logistischen Regression nach den folgenden Gleichungen zu den jeweiligen deterministischen Komponenten, wobei die β-Koeffizienten die Endpunktgewichte liefern:

(i) $V(A) = \alpha + \beta_1 X_1^A + ... + \beta_n X_n^A$

(ii) $V(B) = \alpha + \beta_1 X_1^B + ... + \beta_n X_n^B$

und

(iii) $\ln {}^{p(A)}/_{p(B)} = \beta_1(X_1^A - X_1^B) + ... + \beta_n(X_n^A - X_n^B)$

Theoretische Verortung: ökonomische Nutzentheorie (CA) versus deskriptive Theorie (AHP)

Zwar fußen beide Verfahren auf den Axiomen der Nutzentheorie, aber beim AHP wird aufgrund der Schwierigkeiten präziser Präferenzbeurteilung durch Anwender auf restriktive Annahmen der Konsistenz und Transitivität der Präferenzäußerungen verzichtet. Somit werden wegen der weniger strengen Annahmen hinsichtlich der Präferenzstruktur auch teils irrationale nicht als falsche Entscheidungen identifiziert. Das AHP-Verfahren erfüllt somit wesentliche Anforderungen der ökonomischen Nutzentheorie nicht und gilt nutzentheoretisch als nicht fundiert. Saaty (1990) bezeichnet die dem AHP zugrunde

liegenden Theorie als *deskriptive* - von der *normativen Nutzentheorie* abzugrenzende - Theorie.

Methodische Gütekriterien CA versus AHP

Bei den aus der Sozialempirie, Psychometrie und Biometrie etablierten methodischen Gütekriterien wie Augenscheinvalidität, interne Validität, Konvergenzvalidität, prädiktive Validität und Modellpassung schneiden die beiden Methoden zwar in den jeweiligen Studien unterschiedlich ab, ohne dass sich aber dadurch ein Trendvorteil für eine Methode ergibt. Lediglich bei verschiedenen Dominanztests, bei denen die Übereinstimmung des Ergebnisses in Relation zur von den Befragten vorab als dominierend bezeichneten Alternative gemessen wird, scheint die CA etwas unterlegen zu sein. CA kann bezogen auf a priori Annahmen somit zu widersprüchlicheren Ergebnissen führen. Dominanztest sind ein Kriterium für die prädiktive Validität. Beispielhaft aus der hierzu vorhandenen Literatur seien zwei Studien zitiert (Mulye 1998; Scholl et al. 2005). Triantafyllou (2000) definiert allerdings in seiner Arbeit zwei Kriterien zur Bewertung der methodischen Güte: 1. Akkurate Methoden in multidimensionalen Entscheidungsfindungen sollen auch in eindimensionalen Fragestellungen erfolgreiche Lösungen liefern. Dies wird in den meisten Fällen von AHP (als gewichtetes Summenmodell) erfüllt. 2. Die Reihenfolge der Alternativen hin zur besten soll nicht geändert werden, wenn eine Alternative (nicht die beste) durch eine andere ersetzt wird, wobei die Attributsgewichte unverändert bleiben.

Hier tritt allerdings das Phänomen des Rank Reversal bei AHP auf, also der Nichteinhaltung genau dieses zweiten Gütekriteriums. Als Ursache des Rank Reversal wird die Ermittlung der relativen Gewichtungen mittels der Eigenwertmethode diskutiert. Obwohl das AHP-Verfahren auf den Axiomen der Nutzentheorie basiert, wird aufgrund der Schwierigkeiten einer präzisen Präferenzbeurteilung durch die Anwender auf die restriktiven Annahmen der Konsistenz und Transitivität der Präferenzäußerungen verzichtet. Somit werden wegen der weniger strengen Annahmen hinsichtlich der Präferenzstruktur auch teils irrationale nicht als falsche Entscheidungen identifiziert (Saaty & Vargas 2001). Das AHP-Verfahren erfüllt somit wesentliche Anforderungen der ökonomischen Nutzentheorie nicht und gilt nutzentheoretisch als nicht fundiert. Allerdings kann dieser Argumentation von Seiten der Verhaltensökonomie auch entgegengehalten werden, dass die auf den Annahmen des Homo economicus basierende Nutzentheorie einer nutzenmaximierend agierenden, der ökonomischen Rationalität folgenden individuellen Verhaltensweise nach dem Standardmodel der Ökonomie nicht dem realen menschlichen Verhalten in ökonomischen Entscheidungen entspricht (Tversky/Kahneman 1974).

Bezogen auf die *Rationalität* als *komplexe Verhaltenstheorie* leitet der Nobelpreisträger McFadden (1998) spezifisch für die Aspekte *Präferenzen, Perzeptionen* und *Prozess* über entsprechende Experimente her, dass die meisten kognitiven Anomalien aus Fehlern in der *Perzeption* herrühren, die damit zusammenhängen, wie Informationen gespeichert, abgerufen und verarbeitet werden bzw. aus Fehlern im Entscheidungsprozess selber, die zu einer Formulierung von - kognitiv betrachtet - inkonsistenten Entscheidungsproblemen, zumindest wenn man Rationalität in ihrer engen, nutzentheoretisch fundierten Definition wahrnimmt, führen. Shafir und LeBoef (2002) zeigen auch auf, dass Individuen oft Rationalitätsannahmen zu logischen Schlussfolgerungen, Beurteilungen und Wahlmöglichkeiten in ihren Entscheidungen abweichend vom normativen Verhaltensmodell verletzen. Sie unterstützen die mittlerweile in den Sozialwissenschaften aufkommende Kritik an der Rationalität als axiomatische Verhaltensform.

Methodenvergleich CA versus AHP		
	Conjoint Analyse	**AHP**
Voraussetzung	**Dekompositionell** Vorzugsweise Unabhängigkeit der Attribute, arbiträr sind Kombinationen von Attributsleveln (Ausprägungen) möglich	**Kompositionell** Vorzugsweise Unabhängigkeit der Attribute
Eingesetzte Skalen des Inputs	Ordinal oder Intervallskaliert	Semantisch deskriptive 9-Punkte-Verhältnisskala
Skalen des Outputs	Intervallskala	Verhältnisskala
Anwendbarkeit	Bis zu 6 Attributen mit 2-4 Level, Viele Stimuli	Viele Attribute, bis zu 7-8 Attribute bzw. Attributslevel (Auprägungen)
Methodischer Ansatz	Ermittlung der Teilnutzenbeiträge von Eigenschaften und deren Ausprägungen auf Basis empirisch erhobener Gesamtnutzenurteile	Direkte Bewertung der Eigenschaften und deren Ausprägungen mit anschließender Gesamtbewertung
Bewertungsprozess	Ganzheitlich: reale Entscheidungssituation wird simuliert	Paarvergleichsbewertung: Komplexitätsreduktion durch Dekomposition und hierarchische Strukturierung
Methodische Schwächen	Schnelle kognitive Überforderung der Probanden bei Vielzahl von Eigenschaften und deren Ausprägungen	Bewertung mittels Paarvergleichen entspricht nicht der realen Entscheidungssituation, z. T. mangelnde Stabilität der Ergebnisse
Ressourcenaufwand (Patienten, Zeit, etc.)	Patientenzahl bei repräsentativer Stichprobe oft über 100, digital und auf Papier machbar, mindestens 6 Wochen	Patientenzahl bei repräsentativer Stichprobe ab 25, Digital und auf Papier machbar, ca. 2 bis 3 Wochen

Tab. 2: Methodenvergleich: CA versus AHP.

Kognitive Überforderung (CA) versus unreale Entscheidungssituationen (AHP)

Einem fiktiven Beispiel mit fünf Attributen (N=5) und jeweils drei Ausprägungen bzw. Levels (L=3) für ein zu bewertendes Gut mit Alternativen folgend, ergeben sich nach den weiter oben beschriebenen mathematischen Grundlagen beider Verfahren beim AHP n (n-1)/2 = 10 Paarvergleiche (Trade-offs) und bei der CA 5^3 = 125 Szenarien, was bei einem simplifizierenden sogenannten orthogonalem Design, wonach die Annahme gilt, dass keine Interaktionseffekte zwischen den Attributen bestehen und somit keine Interaktionseffekte gemessen werden müssen mit der Konsequenz, dass jede Ausprägung eines Attributs mit jeweils einer Ausprägung eines anderen Attributes in einem Szenario vorkommen, 12 Szenarien pro befragten Patient ergibt. Berücksichtigt man die Trade-offs innerhalb der Szenarien erscheint die kognitive Herausforderung, ja sogar Überforderung bei der CA selbstevident im Vergleich zum AHP. Andererseits wird die Versorgungsrealität allerdings viel besser über solche Szenarien abgebildet, als durch abstrakte paarweise Vergleiche beim AHP. Abbildung 2 zur paarweisen Befragung beim AHP und Abbildung 3 zu den präferierten Szenarien im Rahmen einer CA, beide patientenseitige Endpunktgewichtungen zum Krankheitsbild Hepatitis C betreffend, sollen die jeweiligen Herausforderungen verdeutlichen.

Vergleichende Studien CA versus AHP

Ijzerman et al (2012) vergleichen AHP und CA bei der Bewertung von Alternativen in der Rehabilitation von Schlaganfällen. In den Gewichtungen der betrachteten Attribute dieser Alternativen - wobei aus

methodischen Gründen die Attribute des einen Verfahrens für das andere zusammengelegt werden müssen - kommen sie zum Ergebnis, dass es eine Übereinstimmung der jeweiligen Gewichte bei vergleichbaren Mittelwertstandartabweichungen für drei Attribute gibt (AHP versus CA: klinischer Outcome 0,30 versus 0,33; Einfachheit der Anwendung 0,16 mit beiden Methoden, Komplikationen 0,20 versus 0,19), Abweichungen dagegen bei zwei Attributen (AHP versus CA: Kosmetik und Komfort 0,28 versus 0,05; Einfluss der Behandlung operationalisiert als Dauer und Invasivität 0,06 versus 0,28). Der Vergleich zeigt darüber hinaus Unterschiede im Ressourcenverbrauch und in der kognitiven Beanspruchung der Befragten. Die CA ist aufwendiger und komplizierter.

Andere Vergleiche wiederum, wie die Studie von Ijzerman et al. aus dem Jahre 2008 zur Erhebung von Präferenzen bei Patienten mit neurologischen Störungen, zeitigen relativ ähnliche Attributsgewichte, wobei nicht-medizinische Faktoren wie soziale Akzeptanz und Anwenderfreundlichkeit der jeweiligen unterstützenden Behandlungen eine große Rolle zu spielen scheinen. In dieser Studie unterschieden sich die Gewichte zwischen Patienten und befragten klinischen Experten unabhängig von der angewandten Methode. Ein Phänomen, auf das man auch in anderen Indikationen wie beispielsweise der Depressionsbehandlung trifft (Hummel et al. 2012; Danner et al. 2011). In einer weiteren Studie (Benaïm et al. 2010) verwenden die Autoren CA und AHP zur Gewichtung der 6 Items der London Handicap Scale, eines Erhebungsinstruments der Beeinträchtigung, das bei der Schlaganfallprävention oft Anwendung findet. Mit beiden Methoden wird die physische Unabhängigkeit am höchsten gewichtet, Unterschiede zeigen sich jedoch in der Reihenfolge der anderen Items. Bei Anwendung der sogenannten „10 binary choices“ Methode nach Schoenemaker & Waid (1982), wonach 10 zufällige fiktive Patientenprofilpaare mit ähnlichen Summenscores gewählt werden und die Konkordanz zwischen schlechteren Profilen mit AHP & CA Ergebnissen verglichen wird (ein Kappa-Wert > 0,40 führt zu einer guten Übereinstimmung, ein Kappa-Wert ≤ 0 zeigt eine sehr schlechte Übereinstimmung), sind bei der CA in 3 von 10 Fällen gute Übereinstimmungen erzielt worden, in 1 eine moderate und in 6 ein schlechte Übereinstimmung, wohingegen beim AHP in 6 von 10 Fällen eine gute und in jeweils 2 eine moderate und eine schlechte Übereinstimmung erzielt wurden.

Verwertbarkeit von CA und AHP für das IQWiG

In ihrem Diskussionspapier zum Methodenvergleich bei der Gewichtung multipler patientenrelevanter Endpunkte entwickeln Neidhardt et al. (2012) einen Kriterienkatalog. Dieser orientiert sich an (i) den gesetzlichen Vorgaben, (ii) den Anforderungen aus der IQWiG-Methodik (Effizienzgrenzkonzept) und (iii) den allgemeinen Anforderungen an das wissenschaftliche Vorgehen. Unter den gesetzlichen Vorgaben subsumieren die Autoren die Transparenz und ein international etabliertes Verfahren der Gesundheitsökonomie. Zu den Anforderungen aus der IQWiG-Methodik zählen sie die approximative Kardinalität des aggregierten Nutzenmaßes, die Flexibilität und die Möglichkeit zur Aggregation

Literatur

Benaïm, C./Perennou, D. A./Pelissier, J. A./Daures, J. P. (2010): Using an analytical hierarchy process (AHP) for weighting items of a measurement scale: A pilot study. In: Revue d'Épidémiologie et de Santé Publique 2010, 58: 59–63

Danner, M./Hummel, J. M./Volz, F./van Manen, J. G./Wiegard, B./Dintsios, C. M./Bastian, H./Gerber, A./Ijzerman, M. J. (2011): Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences. Int J Technol Assess Health Care 2011, 27, 4: 369-75

De Bekker-Grob, E./Ryan, M./Gerard K. (2012): Discrete Choice Experiments in Health Economics: A review of the literature. In: Health Econ 2012, 21: 145–172

EMA (2010): Benefit-risk methodology project. Work package 2 report: Applicability of current tools and processes for regulatory benefit-risk assessment. In: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2010/10/WC500097750.pdf (abgerufen am 16.01.2013)

FDA (2012): Guidance for Industry and Food and Drug Administration Staff - Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approvals and De Novo Classifications. In: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM296379.pdf (abgerufen am 16.01.2013)

Hummel, M. J./Volz, F./van Manen, J. G./Danner, M./Dintsios, C. M./Ijzerman, M. J./Gerber, A. (2012): Using the analytic hierarchy process to elicit patient preferences: prioritizing multiple outcome measures of antidepressant drug treatment. In: Patient 2012, 5, 4:225-37

IJzerman, M. J./van Til, J. A./Bridges, J. (2012): A Comparison of Analytic Hierarchy Process and Conjoint Analysis Methods in Assessing Treatment Alternatives for Stroke Rehabilitation. In: Patient 2012, 5, 1: 45-56

IJzerman, M. J./van Til, J. A./Snoek, G. J. (2008): Comparison of Two Multi-Criteria Decision Techniques for Eliciting Treatment Preferences in People with Neurological Disorders. In: Patient 2008, 1, 4: 265-272

IQWiG (2011): Allgemeine MethodenVersion 4.0. In: https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf (abgerufen am 16.01.2013)

Lancaster, K. (1966): A New Approach to Consumer Theory. In: Journal of Political Economy 1966, 74:132-57

Lootsma, F.A./Schuijt, H. (1997): The multiplicative AHP, SMART and ELECTRE in acommon context. In: Journal of Multi-Criteria Decision Analysis 1997, 6: 185-196

McFadden, D. (1999): Rationality for Economists? In: Journal of Risk and Uncertainty 1999, 19, 1-3: 73-105

Mulye, R. (1998): An empirical comparison of three variants of the AHP and two variants of conjoint analysis. In: Journal of Behavioral Decision Making 1998, 11, 4: 263-280

Neidhardt, K./Wasmuth, T./Schmid, A. (2012): Die Gewichtung multipler patientenrelevanter Endpunkte - Ein methodischer Vergleich von Conjoint Analyse und Analytic Hierarchy Process unter Berücksichtigung des Effizienzgrenzenkonzepts des IQWiG. In: Diskussionspapier 02-12 Februar 2012. Universität Bayreuth, Wirtschaftswissenschaftliche Diskussionspapiere.

Saaty, T. L./Vargas, L. G. (2001): Models, methods, concepts & applications of the analytic hierarchy process. In International series in operations research & management science 2001, 34, Boston: Kluwer Academic Publishers

Saaty, T.L. (1990): An exposition of the AHP in reply to the paper "Remarks on the analytic hierarchy process". In: Management Science 1990, 36, 3: 259-268

Saaty, TL. (1989): Group decision making and the AHP. In: Golden, B. L./Wasil, E. A./Harker, P. T. (Hrsg.) (1989). The Analytic Hierarchy Process: Applications and Studies. Berlin; Springer

Schoenemaker, P./Waid, C. (1982): An experimental comparison of different approaches to determining weights in additive utility models. In: Management science 1982, 28: 182-196

Scholl, A./Manthey, L./Helm, R./Steiner, M. (2005): Solving multiattribute design problems with analytic hierarchy process and conjoint analysis: An empirical comparison. In: European Journal of Operational Research 2005, 164, 3: 760-777

Shafir, E./LeBoeuf, R. A. (2002): Rationality. In: Annu. Rev. Psychol. 2002, 53: 491–517

Triantaphyllou, E. (2000): Multi-Criteria Decision Making Methods: A Comparative Study. Boston: Kluwer Academic Publishers

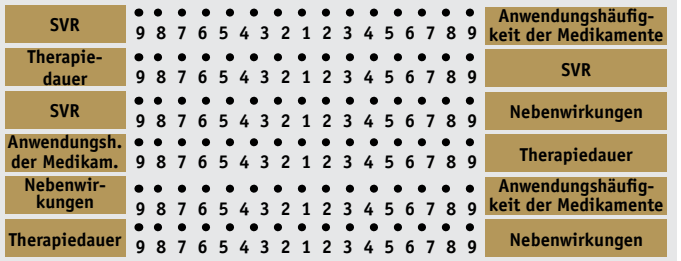
Tversky, A./Kahneman, D. (1974): Judgment under uncertainty: heuristics and biases. In: Science 1074, 185:1124-1131

Wallenius, J.7Dyer, J. S./Fishburn, P. C./Steuer, R. E./Zionts, S./Deb, K. (2008): Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. In: Management Science 2008, 54, 7: 1336-1349

Zanakis, S. H./Solomon, A./ Wishart, N./ Dublish, S (1998): Multi-attribute decision making: a simulation comparison of select methods. In: European Journal of Operational Research 1998, 107: 507-529

Gewichtung von Aspekten der Hepatitisbehandlung

Jeder Aspekt soll im Folgenden von Ihnen als behandelnder Arzt gewichtet werden, d.h. wenn Sie an die Zielsetzung der Hepatitis-Behandlung ihrer Patienten denken, welchem Aspekt würden Sie eine größere Bedeutung beimessen? „Welcher Aspekt der Hepatitis-Behandlung ist Ihnen wichtiger, und um wie viel wichtiger?“



©MVF

Abb. 2: Paarweise Vergleiche zur Endpunktgewichtung bei Hepatitis C; Legende: 1 = gleich wichtig, 2= liegt dazwischen, 3 = etwas wichtiger, 4 = liegt dazwischen, 5 = viel wichtiger, 6 = liegt dazwischen, 7 = sehr viel wichtiger, 8 = liegt dazwischen, 9 = extrem wichtiger.

CA-Szenarien zur Endpunktgewichtung bei Hepatitis C

	Therapie A	Therapie B
Dauer der antiviralen Therapie	24 Wochen	24 Wochen
Das Medikament (Interferon) muss gespritzt werden	1x alle 14 Tage	1x pro Woche
Anhalten grippeähnlicher Symptome nach der Injektion	an 3 Tagen nach der Inj.	an 1 Tag nach der Inj.
Die Wahrscheinlichkeit, dass Magen-Darm-Beschwerden auftreten, liegt bei	25 von 100 Pers. (25 %)	45 von 100 Pers. (45 %)
Die Wahrscheinlichkeit, dass psychische Nebenwirkungen auftreten, liegt bei	55 von 100 Pers. (55 %)	45 von 100 Pers.(45 %)
Die Wahrscheinlichkeit, dass Hautprobleme und/oder Haarausfall auftreten, liegt bei	55 von 100 Pers. (55 %)	45 von 100 Pers.(45 %)
Die Wahrscheinlichkeit der anhaltenden Virusfreiheit 6 Monate nach Therapieende liegt bei	55 von 100 Pers. (55 %)	45 von 100 Pers. (45 %)
Bitte entscheiden Sie sich für eine Therapie (kreuzen Sie an)	☐	☐

Tab. 3: CA-Szenarien zur Endpunktgewichtung bei Hepatitis C.

von Präferenzen. Als allgemeine Anforderungen an das wissenschaftliche Vorgehen definieren sie die nutzentheoretische Fundierung, die Erfüllung von vorbestimmten Gütekriterien, die Wirtschaftlichkeit und Praktikabilität des Ansatzes und die Vollständigkeit der Bewertung mittels der eingesetzten Methoden. Die Bewertung der Methoden anhand dieses Kriterienkatalogs schlussfolgern die Autoren, konnte zeigen, dass weder CA noch AHP alle Anforderungen erfüllen, die an eine Methode zur Endpunktgewichtung im Rahmen des Effizienzgrenzenansatzes unter Berücksichtigung der gesetzlichen Vorgaben, der IQWiG-Methodik und der allgemeinen Anforderungen an wissenschaftliches Vorgehen zu stellen sind und sie empfehlen die Durchführung weiterer empirischer Vergleichsstudien, die beide Methoden gezielt im Hinblick auf die Fragestellung der Verwertbarkeit zur Endpunktgewichtung im Rahmen der IQWiG-Methodik überprüfen. Wenn man allerdings die bereits vorliegenden und auch hier vorgestellten Vergleichsarbeiten hinsichtlich des vorgestellten Kriterienkatalogs analysiert, gelangt man zur selben Schlussfolgerung. Letztlich müssen die Anwender die Kriterien des vorgestellten Kriterienkatalogs anhand ihrer eigenen Präferenzen priorisieren, um sich für das eine oder andere Verfahren zu entscheiden (Tab. 3).

Patient-weighted endpoints: Conjoint Analysis vs. Analytic Hierarchy Process

Both techniques, Conjoint Analysis (CA) and Analytic Hierarchy Process (AHP), are used to elicit patient preferences (in this case to prioritize and weight endpoints). Beyond their theoretical basis (economical utility theory versus descriptive theory) they differ significantly concerning the underpinned model (additive partial utility model versus versus weighted additive utility model), the methodological approach (decompositional versus compositional), the assessment process (holistic versus pair-wise comparisons), methodological weaknesses (excessive demand of cognition versus unreal decision scenarios) and the resource use (extensive versus low). CA may lead in relation to a priori assumptions to conflicting results. The AHP's drawback is a possible rank reversal. Both techniques are generating in comparing studies quite similar results and show depending on the research question they are used pros and cons respectively.

Keywords

Multicriterial decision analysis, Conjoint Analysis, Analytical Hierarchy Process, Patients' preferences, Endpoint weighting, Benefit-Risk trade off, Benefit aggregation

Ausblick

Multikriterielle Entscheidungsverfahren zur Patientenpräferenzerhebung bei Endpunktgewichtung gewinnen immer mehr an Bedeutung in Gesundheitssystemen. In Deutschland sind sie für die gesundheitsökonomische Evaluation nach dem vom IQWiG vorgeschlagenen Effizienzgrenzenansatz eine Prämisse, um endpunktbezogenen Nutzen zu aggregieren. Das IQWiG nennt in seinem aktuellen Methodenpapier (Version 4.0) CA und AHP auch als eine Möglichkeit, endpunktbezogene Nutzen- und Schadensaspekte gegenüberzustellen (IQWiG 2011).

Die FDA (2012) sieht in ihrer Guidance for Industry and FDA Administration Staff 03/2012 unter der Rubrik „Factors to Consider When Making Benefit-Risk Determinations“ Patientenpräferenzen hinsichtlich der Nutzen- und Schadensaspekte von medizinischen Anwendungen als nützliche Informationsquelle. Aber auch die EMA (2010) hat den Stellenwert der Ermittlung von Patientenpräferenzen durch multikriterielle Entscheidungsanalyseverfahren zur Nutzen-Schaden-Abwägung in ihrem Benefit-Risk Methodology Project erkannt. Es bleibt abzuwarten, ob auch Entscheidungsträger wie der G-BA sich Methoden wie die CA und das AHP bei ihren Beschlussfassungen zum Zusatznutzen von neuen Arzneimitteln im Rahmen der frühen Nutzenbewertung für die Saldierung von Nutzen- und Schadensaspekten anstelle intuitiver Heuristiken bedienen werden. Patientenrelevante Endpunkte, die den Nutzen und Schaden für die betroffenen Patienten abbilden, sollten zur Herleitung einer Gesamtaussage hinsichtlich eines vorliegenden Zusatznutzens in Zeiten von Shared-decision making auf der Arzt-Patienten-Entscheidungsebene von den betroffenen Patienten auch auf Systemebene gewichtet werden. Dies verlangt auch, dass das System dem mündigen Patienten Vertrauen schenkt, und zwar nicht nur auf der rein rhetorisch-verbalen Ebene, sondern auch im Rahmen der unmittelbar die betroffenen Patienten tangierenden Entscheidungsfindung. <<

Dr. rer. biol. hum. Ch.-Markos Dintsios, MPH

ist seit 2011 Senior Referent für Gesundheitsökonomie beim Verband der forschenden Arzneimittelhersteller (vfa) in Berlin. Der Pharmazeut, Wirtschaftswissenschaftler und Master in Public Health war 1994-1997 Patientenvertreter am Klinikum Nürnberg, bis 2005 wissenschaftlicher Mitarbeiter am Institut für Epidemiologie, Sozialmedizin & Gesundheitssystemforschung der Medizinischen Hochschule Hannover und von 2005-2011 beim IQWiG tätig. Kontakt: c.dintsios@vfa.de



Dr. med. Thomas Stark

Surrogate und Endpunkte beim Prostatakarzinom

Zusammenfassung

Mit der Verfügbarkeit neuer Therapiemöglichkeiten hat sich die Prognose für Patienten mit metastasiertem kastrationsresistenten Prostatakarzinoms (mCRPC) deutlich verbessert. Aufgrund der Entwicklungszyklen ergibt sich für zukünftige Neuentwicklungen von Medikamenten die Notwendigkeit, Surrogatendpunkte zu definieren, die prädiktiv für das Gesamtüberleben und ein Therapienansprechen sind und gleichzeitig von den Gesundheitsbehörden anerkannt werden.

Schlüsselwörter

Prostatakrebs, metastasiert, kastrationsresistent, hormonresistent, Surrogat, Endpunkt

Das Prostata-Karzinom ist die häufigste Krebserkrankung des Mannes. Es wird geschätzt, dass in Deutschland im Jahr 2012 etwa 68.000 Männer an einem Prostatakarzinom erkrankten. Wenngleich die Sterblichkeitsrate seit 2003 sinkt, so sind nach Angaben des Robert Koch-Instituts im Jahr 2008 immer noch rund 11.000 Männer an Prostatakrebs verstorben [Robert Koch- Institut 2012]. Dabei wird deutlich, dass es sich bei dem Prostatakarzinom um eine Erkrankung handelt, die auch eine große gesundheitsökonomische Bedeutung hat. Mit der Entwicklung neuer wirksamer Medikamente hat sich das Überleben von Patienten mit metastasiertem kastrationsresistentem Prostatakarzinom deutlich verbessert. So zeigte sich in einer kürzlich vom Institute of Cancer Research vorgestellten Untersuchung in einem Patientenkollektiv von 442 Patienten ein medianes Überleben von 31 Monaten, was wesentlich länger ist als das Gesamtüberleben, das in früheren Studien und anhand von Nomogrammen abgeleitet werden konnte, das bei 18 Monaten (Halabi) bzw. 21 Monaten (Smaletz) lag (Abb. 1 [Omlin et al. 2013]). Veranschaulicht man sich außerdem das geschätzte Überleben von 6 bis 10 Monaten dieser Patienten in der therapeutischen Ära vor der Zulassung von Docetaxel im Jahr 2005, so sieht man, dass mittlerweile mehr als eine Verdoppelung des Überlebens erreicht werden konnte.

>> Auch wenn in den letzten beiden Jahren ein enormer Entwicklungssprung in der medikamentösen Behandlung des fortgeschrittenen Prostatakarzinoms mit der Zulassung von Cabazitaxel und Abirateronacetat stattgefunden hat, so sind derzeit über 600 Substanzen in der weiteren Entwicklung - davon stehen 11 Medikamente zur Behandlung des fortgeschrittenen Prostatakarzinoms bereits kurz vor der Zulassung bzw, in Zulassungsstudien. Weitere Medikamente mit neuen Wirkmechanismen geben Anlass zur Hoffnung, mittelfristig einen Paradigmenwechsel in der Behandlung des fortgeschrittenen Prostatakarzinoms zu erleben, so dass diese Erkrankung nicht mehr häufig tödlich, sondern nahezu regelmäßig chronisch verlaufen könnte.

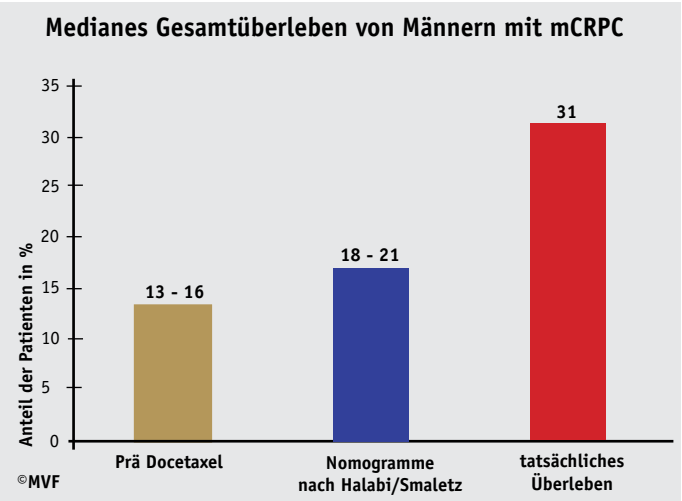


Abb. 1: gegenwärtiges medianes Gesamtüberleben von Männern mit mCRPC im Vergleich zu Prognosen nach Nomogrammen und historischen Kohorten. Abbildung erstellt nach Omlin et al. 2013.

Dabei wird aber auch deutlich, dass man sich mit der besseren Prognose der Patienten von dem bisherigen Goldstandard der Endpunkte des mCRPC, dem Gesamtüberleben, lösen muss, wenn der Weg zur weiteren Entwicklung neuer Medikamente nicht verstellt werden soll.

Die Europäische Zulassungsbehörde trägt in ihren Empfehlungen zu klinischen Studien diesem Umstand Rechnung und akzeptiert neben dem Gesamtüberleben auch die Endpunkte „Progressionsfreies Überleben“, „Symptomkontrolle“ bzw. „Zeit bis zum symptomatischen Progress“ als primäre Endpunkte für konfirmatorische Phase-III-Studien für antineoplastische Medikamente [EMA 2011]. Empfohlen wird allerdings, Gesamtüberleben zumindest als sekundären Endpunkt zu erfassen, um so nachteilige Effekte der Therapie, z.B. auf die Wirksamkeit von Folgetherapien, abschätzen zu können. Unabhängig davon, bleibt Gesamtüberleben natürlich ein patientenrelevanter Endpunkt.

Folgetherapie als methodisches Problem zur Beurteilung des Gesamtüberlebens

Das Beispiel des Nierenzellkarzinoms zeigt, dass ein Beharren auf Daten zur Verbesserung des Gesamtüberlebens eine Zulassung dieser Medikamente zumindest deutlich verzögert hätte: Bei den meisten der zugelassenen Target-Therapien für das metastasierte Nierenzellkarzinom (mRCC) zeigte sich zwar ein signifikanter Vorteil im progressionsfreien Überleben, nicht jedoch für das Gesamtüberleben. Vor Zulassung der ersten Targettherapien hatte sich das mRCC als weitgehend resistent gegenüber allen heute verfügbaren Chemotherapien gezeigt, auch die Strahlentherapie erwies sich als wenig effektiv.

Bis zur Einführung der zielgerichteten Medikamente war die Immuntherapie die Standardbehandlung. Bei diesen Therapien mit Interferon-Alpha (IFN--) und Interleukin-2 lagen die Ansprechraten unter 10 bis 20 Prozent, daneben ist der Einsatz der Immunmodulatoren mit erheblichen Nebenwirkungen verbunden. Das mediane Gesamtüberleben unter Immuntherapien lag bei 5 bis 14 Monaten, es wurden maximal 30 Monate für eine gute Risikogruppe erreicht, dies allerdings unter sehr hoher Toxizität. Im klinischen Alltag ging man bei mRCC-Patienten vor 2006 von einer medianen Überlebenszeit von 6 bis 12 Monaten aus. In einer kürzlich publizierten retrospektiven Analyse von mRCC-Patienten, die im Rahmen einer Zulassungsstudie mit Everolimus behandelt worden waren, zeigte sich im weiteren Verlauf ein medianes Gesamtüberleben von 40,5 Monaten.

An diesem Beispiel wird deutlich, dass alternative Endpunkte wie das radiographisch beurteilte progressionsfreie Überleben ein sinn-

volles Surrogat für das Gesamtüberleben darstellen kann, auch wenn eine Validierung dieses Endpunktes als Surrogatendpunkt für einzelne Therapieoptionen teilweise noch erforderlich ist. Auch wenn in den jeweiligen Zulassungsstudien (mit Ausnahme von Temsirolimus in einer prognostisch schlechten Prognosegruppe) der Nachweis einer Verbesserung des Gesamtüberlebens nicht gelang, so zeigt sich zudem, dass die Daten in der Realität schließlich anders und teilweise auch deutlich besser ausfallen können als im Rahmen einer Zulassungsstudie.

Alternative Endpunkte beim Prostatakarzinom bislang nicht etabliert

Beim Prostatakarzinom ergibt sich in einzelnen Krankheitsstadien allerdings das Problem, dass sich die Tumormasse im metastasierten Stadium oft nicht gut bildgebend darstellen und damit quantifizieren lässt. Daher empfiehlt die EMA wie auch die Prostate Cancer Working Group, für das progressionsfreie Überleben nur neu aufgetretene Knochenläsionen im Knochenszintigramm zu werten und eine objektive Bewertung messbarer Läsionen anhand modifizierter RECIST-Kriterien durchzuführen. Insbesondere in den ersten 12 Wochen einer Behandlung des Prostatakarzinoms sollten dabei paradoxe Phänomene wie das des „Bone-Flares“ berücksichtigt werden. Dabei kann es durch die vermehrte Stoffwechselaktivität, die durch Zellzerfall oder Remodeling-Prozesse im Knochen entstehen, im Skelettszintigramm vorübergehend fälschlicherweise so erscheinen, als wäre es zu einem Progess der Knochenmetastasen gekommen. Bei der Wahl alternativer Endpunkte für das fortgeschrittenen bzw. metastasierte Prostatakarzinoms sollte außerdem berücksichtigt werden, dass dieser Tumor eine hohe biologische Variabilität aufweist. Die Biologie des Tumors verändert sich im Laufe der Erkrankung sowohl aufgrund seiner genetischen Instabilität, als auch aufgrund veränderter Umgebungsfaktoren wie etwa einer durchgeführten antineoplastischen Therapie.

Es erstaunt daher nicht, dass das prostataspezifische Antigen (PSA), in seiner diagnostischen Wertigkeit Veränderungen im Verlauf der Erkrankung erfährt. Ein erneut ansteigendes PSA nach vorangegangener operativer Entfernung der Prostata (mit der Intention, die Erkrankung zu heilen) hat eine sehr hohe prognostische Wertigkeit hinsichtlich der Vorhersage eines Rezidives oder Progresses der Erkrankung. Im weiter fortgeschrittenen Verlauf, also z.B. bei Patienten mit einem metastasierten, kastrationsresistenten Prostatakarzinom, ist die prognostische Wertigkeit des PSA-Wertes zwar teilweise noch hinsichtlich des weiteren Überlebens des Patienten gegeben, jedoch nicht mehr zwingend zur Beurteilung eines Therapieansprechens.

Validierung neuer Endpunkte benötigt den Zusammenschluss von Wissenschaft, Industrie und Behörden

Die Geschichte der Entwicklung neuer Medikamente zur Therapie der HIV-Infektion kann ggf. als Vorbild dienen, den richtigen Weg auch in der Indikation des fortgeschrittenen Prostatakarzinoms einzuschlagen. Wohl kaum eine medizinische Indikation hat in den letzten Jahrzehnten einen derartigen Innovationsschub, und damit einhergehend, eine deutliche Verbesserung der Behandlung erfahren wie HIV und AIDS. Eine solche Dynamik mit der Entwicklung von mehr als 20 verschiedenen Medikamenten innerhalb kurzer Zeit (1987 - 1994 vier Medikamente, 1995 – 2006 20 Medikamente) war unter anderem aber auch nur deshalb möglich, weil neben der AIDS-assoziierten Mortalität ein neues Surrogat zur Beurteilung einer erfolgreichen Behandlung eingeführt wurde, nämlich der Abfall der HIV-Viruslast unter die Nachweisgrenze. Damit konnte die Dauer klinischer Studien deutlich verkürzt werden auf 24 Wochen bei beschleunigtem, bzw. 48 Wochen bei normalem Zulassungsverfahren, und damit auch die Zeit bis zur Verfügbarkeit dieser Medikamente in der täglichen Praxis reduziert werden.

Grundlage für diesen Wandel war eine enge Zusammenarbeit von Behörden, Wissenschaftlern und Industrie.

Um der Vielzahl neuer Wirkmechanismen in der Behandlung von Krebserkrankungen gerecht zu werden, stellte die EMA bereits 2010 in einem Positionspapier fest, dass Änderungen in den Empfehlungen zur Evaluierung der Wirkung von Krebsmedikamenten notwendig sind. Neu ist in der im Dezember 2011 zur Diskussion gestellten Leitlinie die Berücksichtigung der Tatsache, dass die Wirksamkeit einiger Medikamente nicht durch eine Bewertung der Krankheitsprogression und anderer Endpunkte anhand bisher etablierter Methoden nachgewiesen werden kann, und daher neue Methoden und Endpunkte in Betracht gezogen werden müssen.

Prädiktive Marker sind für die weitere Entwicklung neuer Medikamente notwendig

Unabhängig von der Diskussion über sinnvolle Endpunkte zur Anwendung in klinischen Studien ist jedoch auch eine bessere Stratifizierung der Patienten vor Studieneinschluss anhand geeigneter Paramater erforderlich, da nur bei Einschluss geeigneter Patienten eine Verbesserung der entsprechenden Endpunkte erzielt werden kann, und Verwässerungseffekte vermieden werden. Dafür sind neben prognostischen Parametern auch prädiktive Parameter notwendig, die zeigen,

Literatur

Krebs in Deutschland 2007/2008. 8. Ausgabe. Robert Koch-Institut (Hrsg) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg). Berlin, 2012.

Omlin, A./Pezaro, C./Mukherji, D./Mulick, A./Cassidy, S./Sandhu, S./Bianchini, D./Olmos, D./Ferraldeschi, R./Maier, G./Thompson, E./Parker, C./Attard, G./de Bono, J (2013): Improved Survival in a Cohort of Trial Participants with Metastatic Castration-resistant Prostate Cancer Demonstrates the Need for Updated Prognostic Nomograms.
In: European Urology Published online 8 January, 2013, http://www.europeanurology.com/article/S0302-2838(12)01559-X/abstract abgerufen 09.01.2013.

EMA Oncology Working Party (2011): Guideline on the evaluation of anticancer medicinal products in man, Appendix I. Doc. Ref. EMA/CHMP/27994/2008 2011.
In :http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/12/WC505001199.pdf abgerufen 09.01.2013..

welche Patienten auf welche Therapie ansprechen bzw. dafür geeignet sind. Während es jedoch bei manchen Tumoren bereits Empfehlungen zur Nutzung prädiktiver Faktoren für die Therapie und die Anwendung in klinischen Studien gibt, wie zum Beispiel die EGFR-Mutation beim Bronchialkarzinom, gibt es beim mCRPC noch keine validierten prädiktiven Marker. Umso mehr sollte es Ziel weiterer Phase III-Studien sein, solche potentiellen Marker zu evaluieren und zu validieren.

Ziel gemeinsamer Anstrengungen sollte daher sein:

- bessere tumor- bzw. krankheitsspezifische Auswahlkritieren für Studienpatienten zu definieren, die eine geeignete Patientenstratifizierung und damit einen größeren „statistischen“ und tatsächlichen klinischen Benefit für Patienten ermöglichen,
- validierte Surrogatmarker zu finden, und
- damit auch alternative Endpunkte zum Gesamtüberleben zu identifizieren.

Wichtig sollte neben dem Kriterium der Objektivierbarkeit eines Surrogatendpunktes stets auch das Befinden des Patienten sein. Insbesondere im Bereich des fortgeschrittenen Prostatakarzinoms, in dem die starke Einschränkung der Lebensqualität aufgrund des Tumorleidens im Vordergrund steht, sollten patientenrelevante Endpunkte wie die Lebensqualität mehr im Vordergrund stehen. Die aktuelle Fokussierung auf das Gesamtüberleben – so wichtig und relevant dieser Endpunkt zumindest teilweise auch sein mag – sollte in Konsens mit Patienten, wissenschaftlicher Community und Industrie hinsichtlich einer flexiblen Gewichtung der Endpunkte auf die jeweiligen Bedürfnisse eines klinischen Settings angepasst werden.

Zusammenfassung

Um zukünftige Medikamentenentwicklungen im Bereich des mCRPC zu ermöglichen, ist das Gesamtüberleben als alleiniger primärer Endpunkt nicht länger aufrechtzuerhalten. Während wir weiter mit dazu beitragen müssen, die Entwicklung neuer Surrogatendpunkte voranzutreiben, müssen aber jetzt schon vorhandene Surrogatendpunkte pragmatisch und patientenorientiert genutzt werden, um weitere Fortschritte in der Behandlung des Prostatakarzinoms zu ermöglichen. Hier sind gemeinsame Anstrengungen von Industrie, wissenschaftlicher Community und Behörden notwendig, um im Sinne der Patienten zeitnah validierte Surrogatendpunkte zu etablieren, und für klinische Studien ein flexibles Bewertungssystem einzuführen, das der jeweiligen Biologie des Tumors und dem entsprechenden Surrogatendpunkt gerecht wird. <<

Surrogates and Endpoints in Prostate Cancer

The prognosis of patients with metastatic castration resistant prostate cancer (mCRPC) has significantly improved with the availability of new treatment options. Lifecycle management of new drug developments necessitates the definition of surrogate endpoints which are predictive for overall survival and treatment response and which are acknowledged by health authorities.

Keywords

Prostate cancer, metastatic, castration resistant, hormone resistant, surrogate, endpoint

Autorenerklärung

Dr. med. Thomas Stark ist Mitarbeiter des Unternehmens Janssen-Cilag GmbH in Neuss. Eine externe finanzielle und/oder inhaltliche Unterstützung erfolgte nicht.

Dr. med. Thomas Stark
ist Director Medical & Scientific Affairs der Janssen-Cilag Germany Pharmaceutical Companies of Johnson & Johnson. 1993 - 1999: Arzt, Zentrum der Inneren Medizin, Medizinische Klinik II, Hämatologie und Onkologie, Infektiologie (Prof. Hölzer, Prof. Stille) an der Universitätsklinik Frankfurt am Main, Deutschland; 1999 - 2001 Medical Advisor Infectious Diseases GlaxoSmithKline Deutschland; 2001 - 2006 Senior Medical Director, Infectious Diseases Medicine Development Center GlaxoSmithKline, Global Research & Development; 2006 - 2011 Therapeutic Area Leader Infektiologie und Onkologie/Hämatologie der Janssen-Cilag Medical Affairs, Europe, Middle East, Africa. Seit Juli 2011 Director Medical & Scientific Affairs. Kontakt: tstark2@its.jnj.com



VERSORGUNGS monitor FORSCHUNG

Der Endpunkt

Patienten Relevante Endpunkte

Wer definiert, was für den Patienten relevant ist?

Der Fachkongress wurde unterstützt von

arvato BERTELSMANN »HEALTHCARE

icw Technologie für Gesundheit

Vivantes

NOVARTIS

Ihrer Gesundheit zuliebe DeutscheBKK

Pfizer

FRESENIUS KABI

MedicalContact AG

INSIGHT HEALTH