

HTA und aktuelle Herausforderungen: Harmonisierung, Real World Data und Surrogatparameter

HTA – How to tackle pressing challenges: International Harmonization, Real World Data, and Surrogates

Abstract

Health Technology Assessment is one of the standard instruments in support of the decision-making to define the public health services both internationally and in the German health care system. Besides systematic reviews, benefit-harm-analyses, health economic evaluations, and, decision-analytic modelling, especially epidemiological and biometrical questions and methods play a key role. From this perspective discussions on increased European cooperation including calls for wider harmonization are attracting greater interests.

The overall aim is, to present at the GMDS workshop relevant information on this emerging field of harmonization in Europe across similarities and differences in the HTA process. Current developments around the composition of the EU-HTA Network are provided. This network accepts the challenge to define and establish a “Joint Work” across Europe. Special emphasis was placed on the discussion on “Harmonization of HTA: is it a threat or does it mean support?”. Furthermore, methodological discussions and questions are being addressed: “Are Real World Data and Surrogates possible parameters for decision-making or HTA?”

Keywords: HTA, systematic reviews, real world data, policy, technology assessment, surrogate endpoints, biological markers, endpoint determination, thresholds, validation techniques, correlation of data, estimation techniques, biometry, simulation, theoretical study, medicine, evidence-based, decision support techniques

Zusammenfassung

Health Technology Assessment (Gesundheitstechnologiefolgenabschätzung) gehört international und im deutschen Gesundheitswesen zu den Standardinstrumenten der Entscheidungsunterstützung zur Definition des Leistungskatalogs. Neben systematischen Reviews, Nutzen-Schaden-Abwägungen, gesundheitsökonomischen Evaluationen und entscheidungsanalytischen Modellierungen spielen insbesondere epidemiologische und biometrische Fragen und Methoden eine Rolle. Vor diesem Hintergrund geraten die zunehmenden Diskussionen um europäische Zusammenarbeit in den Fokus bis hin zu Rufen nach umfangreicher Harmonisierung.

Die GMDS-Arbeitsgruppe HTA möchte gemeinsam mit den AG/AK Methodik Systematischer Reviews, Gesundheitsökonomie und Medical Decision Making dieses Thema aufgreifen und aktuelle Entwicklungen darstellen. Dies beginnt mit einem europäischen Blick auf den Aufbau des EU-HTA-Netzwerks, das sich der Herausforderung stellt, „Joint Work“ europaweit zu definieren und zu etablieren. Hier rückt die Diskussion um „Harmonisierung von HTA“ in den Vordergrund: Ist dies als Bedrohung oder als Unterstützung zu werten? Im zweiten Teil des Artikels wird vor diesem Hintergrund eine weitere aktuelle methodische Frage aus verschiedenen Perspektiven erörtert: „Sind Real World Data und

Alric Rüter¹
Kirsten H. Herrmann²
Ansgar Hebborn³
Matthias Perleth⁴
Ruth Schwarzer¹
Christoph Schürmann¹
Wiebke Sieben¹
Johanna Gillhaus⁵
Ralf Goertz⁶
Ulli Jeratsch⁶
Friedhelm Leverkus⁵
Wendelin Schramm⁷

1 IQWiG – Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Köln, Deutschland

2 Bristol-Myers Squibb GmbH & Co. KGaA, München, Deutschland

3 F. Hoffmann-La Roche AG, Basel, Schweiz

4 G-BA-Gemeinsamer Bundesausschuss, Berlin, Deutschland

5 Pfizer Deutschland GmbH, Berlin, Deutschland

6 AMS Advanced Medical Services GmbH, München, Deutschland

7 Hochschule Heilbronn, Heilbronn, Deutschland

Surrogate mögliche Parameter im Decision Making oder in der Gesundheitstechnologiefolgenabschätzung?“.

Schlüsselwörter: Gesundheitstechnologiefolgenabschätzung, Health Technology Assessment, Real-World-Daten, Politik, Surrogatendpunkte, biologische Marker, Endpunktbestimmung, Schwellenwerte, Validierungstechniken, Korrelation, Schätztechniken, Biometrie, Simulation, theoretische Studie, evidenzbasierte Medizin, Entscheidungsunterstützungstechniken

Einführung

Mittlerweile gehört Health Technology Assessment (Gesundheitstechnologiefolgenabschätzung) international und im deutschen Gesundheitswesen zu den Standardinstrumenten der Entscheidungsunterstützung zur Definition des Leistungskatalogs. Neben systematischen Reviews, Nutzen-Schaden-Abwägungen, gesundheitsökonomischen Evaluationen und entscheidungsanalytischen Modellierungen spielen insbesondere epidemiologische und biometrische Fragen und Methoden eine Rolle.

In einem HTA-Prozess geht es darum, in einem objektiven (unverzerrt, ohne Bias), transparenten und nachvollziehbaren Prozess Informationen zu einer medizinischen Fragestellung zusammenzuführen, um dann eine Entscheidung zu dieser Fragestellung unter möglichst großer Sicherheit zu treffen. Der erste Schritt ist deshalb vorab die Fragestellung bzgl. Patientengruppe, Intervention, Komparator, Outcome und Studiendesign festzulegen und ein Protokoll zu definieren. Dieses Protokoll demonstriert, dass der Prozess nicht durch Erkenntnisse während des Prozesses gesteuert wurde, oder wenn die Ergebnisse vorliegen. In einer systematischen Recherche wird die gesamte verfügbare Evidenz zur jeweiligen Fragestellung aufgefunden. Diese Literatur wird in einem definierten Verfahren gesichtet und einer kritischen Bewertung anhand des Vier-Augen-Prinzips unterzogen, um systematische Fehler in der Bewertung zu vermeiden. Für die Forschungsfrage relevante Studien werden in einem mehrschrittigen Verfahren herausgefiltert. Daten zu vorab definierten Endpunkten (Mortalität, Morbidität, Lebensqualität...) werden extrahiert und in einer Evidenzsynthese, wenn möglich Meta-Analyse, zusammengeführt und ausgewertet. Ein Ziel dieses Verfahrens ist es u.a. durch Meta-Analysen Patientenzahlen durch das Zusammenführen der Studien zu erhöhen, und damit die statistische Datensicherheit zu erhöhen. In einem HTA-Bericht können mehrere, mindestens zwei, Versorgungsalternativen verglichen werden. In einem letzten Schritt wird eine ökonomische Evaluation durchgeführt. In diesem sehr komplexen Verfahren können unterschiedliche Perspektiven eingenommen werden, u.a. die der Krankenversicherung oder der Gesellschaft. Es werden in verschiedenen Gesundheitssystemen unterschiedliche Methoden und Arten angewandt. Folgende Parameter müssen in einem HTA berücksichtigt werden: Arten und Einsatz der Analysen, Evidenzquelle: Studien- oder Routedaten, Literaturrecherche und Bewertung der ökonomischen Studien, Kostenermittlung, Diskontierung und Sensitivitätsanaly-

sen, Modellierung und Entscheidungsanalysen, Bewertung der ökonomischen Evidenz [1], [2].

Die AG HTA, AG/AK Methoden systematischer Reviews, Gesundheitsökonomie und Medical Decision Making möchten gemeinsam mit den aktuellen Informationen aus dem Bereich um HTAs geben, um weiterführende Diskussionen zu fördern. Dazu erläutern anerkannte Experten in einem Workshop auf der GMDS Tagung in Krefeld 2015 ihre unterschiedlichen und kontroversen Standpunkte zum Thema „Harmonisierung von HTA in der EU“ und beleuchten den Einfluss sogenannter „Real World Data“. Dies wird ergänzt durch Ausführungen zur Surrogatvalidierung. Ziel des Workshops ist es verschiedene Standpunkte aufzuzeigen und zu diskutieren, ohne dabei jedoch ein gemeinsames Fazit oder einen Konsens bilden zu müssen. Die Experten stellen in den jeweiligen Abschnitten ihre Sicht zum Thema dar und verantworten somit den mit ihrem Namen versehenen Abschnitt.

HTA-Netzwerk: Zusammenarbeit oder Harmonisierung

• Alric Rüther

Health Technology Assessment, die Bewertung medizinischer Verfahren und Technologien als Grundlage für Entscheidungen im Gesundheitswesen hat sich etabliert, insbesondere zur Erstattung medizinischer Leistungen. HTA-Institutionen und -Abteilungen schießen wie Pilze aus dem Boden, dreistellig allein in Europa. Das heißt jedoch nicht, dass HTA gleich HTA ist. Die Methoden der Bewertung bis hin zu den Prozessen der Entscheidungsfindung variieren teilweise erheblich. Die EU-Kommission erkannte schon sehr früh das Potenzial von HTA, sichtbar an einer kontinuierlich aufbauenden Reihe von Projekten von EURASSESS 1994 bis hin zum Aufbau des Europäischen Netzwerks HTA, EUnetHTA (siehe: <http://www.eunetha.eu/>). Ein Meilenstein für HTA war die Verabschiedung der Richtlinie zur grenzüberschreitenden Patientenversorgung [3] in 2011 in deren Artikel 15 die EU-Kommission ein kontinuierliches HTA-Netzwerk (HTAN) [4] festschreibt. Dieses HTA-Netzwerk wurde 2013 gegründet. Es hat die Aufgabe sich mit der strategischen HTA-Ausrichtung und -Entwicklung zu befassen. Mitglieder sind Vertreter der Gesundheitsministerien aller Mitgliedsstaaten der EU in Begleitung der entsprechenden HTA-Experten [4]. EUnetHTA gilt als das wissenschaftlich technische Netzwerk der öffentlichen und wissenschaft-

lichen HTA-Organisationen der Mitgliedsstaaten [5]. EUnetHTA ist eng mit HTAN verknüpft, soll jedoch von den Mitgliedstaaten getragen werden. Um das bestmöglich aufzusetzen hat die EU-Kommission mit einer Joint Action 3 (JA3) eine letzte Förderperiode für EUnetHTA eingeleitet. JA3 startet im Juni 2016 im Anschluss an die auslaufende JA2 mit einer Projektdauer von vier Jahren. Ziel ist die Zusammenarbeit europäischer HTA-Organisationen, insbesondere hinsichtlich des Austausches und gemeinsamen Nutzens von HTA-Informationen.

Harmonisierung?

Vor dem Hintergrund der Entwicklungen um HTAN und EUnetHTA wird eine Harmonisierung von HTA in Europa diskutiert. Hier gibt es verschiedene Ansätze. Diese reichen vom Austausch von Studiendaten bis hin zu gemeinsamen EU-weiten HTA-Berichten. Es stellt sich die Frage, ob EU-weite einheitliche HTA-Berichte sinnvoll sind. Vordringlich steht die Annahme, dass der hochkomplexe Prozess der HTA-Erstellung viel effizienter wäre bei einem Bericht für alle Länder. Ein europaweit einheitliches Ergebnis wäre für die Industrie ebenso mit viel geringerem Aufwand zu bewältigen. In diesem Zusammenhang müssen nicht nur die Vielfalt der Gesundheitssysteme in Europa sondern auch kulturelle Unterschiede beachtet werden. Ein HTA-Bericht muss auf den Entscheidungsprozess ausgerichtet sein. Dieser bleibt in all seiner systembezogenen und kulturellen Vielfalt in der uneingeschränkten Verantwortung des jeweiligen Mitgliedsstaates. Aus diesem Grund spricht man im Umfeld von EUnetHTA von HTA-Information. Ein HTA-Report stellt eine umfassende wissenschaftliche Analyse bereit. Eine darüberhinausgehende Empfehlung für eine Gesundheitssystementscheidung wird nicht abgegeben. Dies soll den Mitgliedsstaaten selbst vorbehalten bleiben [6]. Harmonisierung in HTA steht also für wissenschaftliche Zusammenarbeit, Einigung auf methodische Standards und Qualitätsprozesse. EUnetHTA trägt diesen Voraussetzungen Rechnung. Basis ist das HTA-Core Model. Hier soll systematisch das „Herz“ („Core“) eines HTA erarbeitet werden und zwar idealerweise von mehreren HTA-Organisationen gemeinsam (joint report). In der auslaufenden JA2 wurde dies an mehreren Piloten mit unterschiedlichen Ansätzen getestet. Parallel dazu wurden methodische Leitfäden erarbeitet und bereitgestellt, z.B. zu „interner Validität von nicht randomisierten Studien“ oder Literatursuche [7]. Auf den Ergebnissen der JA2 wird mit JA3 eine weitere Joint Action aufbauen. Diese hat zum Ziel die praktischen Applikationen in die Arbeitsprozesse der beteiligten Organisationen effizient einzubringen und ein kontinuierliches Netzwerk zu etablieren. JA3 startet im Juni 2016 mit einem Gesamtbudget von 20 Mio. € (60% Förderung, 40% Eigenanteil) und einer Laufzeit von 4 Jahren. Es sind 77 HTA-Organisationen aus allen EU-Mitgliedsstaaten beteiligt. Zusammenfassend kann gesagt werden, dass es sich bei einer EU-weiten Harmonisierung nicht um eine Adaption oder Angleichung handelt, sondern eher um die Anstrengung den nötigen wissenschaftlichen Aufwand für eine

qualitative Entscheidungsgrundlage auf mehrere Schultern sinnvoll und effektiv zu verteilen. So bleiben mehr Ressourcen zur optimalen Berücksichtigung nationaler Gegebenheiten und Bedürfnisse. Vorsicht ist geboten vor dem Einfluss treibender Kräfte, die über ein Absenken des HTA-Qualitätsniveaus eine Aufwandsreduktion, Erleichterung der Marktzugänge und frühestmöglichen Einsatz am Patienten bewirken wollen. Oberstes Gebot muss die effektive Patientenversorgung mit solidarischem Zugang zu wirksamen Leistungen mit tatsächlichem Nutzen sein.

Health Technology Assessment (HTA) und Real World Data (RWD) aus Sicht eines international operierenden forschenden pharmazeutischen Unternehmens

• Ansgar Hebborn

Die Bedeutung von Daten aus der Routineversorgung („Real World Daten“ – RWD) im Rahmen der systematischen Evaluation von Arzneimitteln im Rahmen von Zulassungs-, Preisfindungs- und Erstattungsentscheidungen nimmt kontinuierlich zu. Dies ist zum einen auf die erleichterte Verfügbarkeit dieser Daten und die Weiterentwicklung von Techniken und Methoden der Datenaufbereitung und -analyse zurückzuführen, zum anderen auf die wachsende Erkenntnis das einzig RWD bestimmte im Zusammenhang mit dem Einsatz von Gesundheitsleistungen stehende Fragen beantworten kann.

RWD als wesentliche Grundlage von kontext-spezifischem HTA

Die künftige Verfügbarkeit als wertvoll anerkannter innovativer Arzneimittel hängt wesentlich davon ab, inwieweit die Präferenzen von Patienten, Ärzten, Kostenträgern und anderen Entscheidungsträgern im Gesundheitswesen von forschenden Pharma-Unternehmen verstanden und in deren Forschungs- und Entwicklungsentscheidungen reflektiert werden können.

In diesem Zusammenhang kommt der systematischen Evaluation von Gesundheitsleistungen („Health Technology Assessment“ – HTA) im spezifischen Kontext nationaler Gesundheitssysteme sowie den auf dieser Basis getroffenen Entscheidungen über deren finanzielle Erstattung und Verfügbarkeit besondere Bedeutung und Signalwirkung für Investitionsentscheidungen forschender Pharma-Unternehmen zu.

Unabdingbare Voraussetzung für eine gesellschaftlich positive Langfristwirkung von HTA auf die Dynamik von Innovationsprozessen ist die adäquate Abbildung von gesellschaftlichen Erwartungen an Leistungen des Gesundheitswesens im Kontext begrenzter ökonomischer Ressourcen. Dies erfordert es, HTA als systematischen multidisziplinären Prozesses durchzuführen, der den

medizinischen, sozialen, ökonomischen, ethischen, legalen und organisatorischen Einfluss von Gesundheitsleistungen auf valide abgebildete Zielparameter für das Gesundheitswesen abschätzt z.B. medizinische Outcomes, rechtliche Durchführbarkeit, gesamtwirtschaftliche Produktivität, fairer Zugang). In diesem Sinne bezeichnet HTA die systematische Evaluation des Einflusses von Gesundheitsleistungen auf relevante Zielparameter im spezifischen Kontext eines (nationalen) Gesundheitssystems.

Dieses Idealbild von HTA findet sich jedoch nur selten in realen Gesundheitssystemen. In vielen Ländern konzentriert sich HTA im Wesentlichen auf die systematische Untersuchung des Einflusses von Gesundheitsleistungen auf einige wenige Entscheidungsparameter mit erwartbar guter Evidenzverfügbarkeit. Wesentliche Evidenz für den Einfluss von Gesundheitsleistungen auf andere, allzu häufig vernachlässigte Entscheidungsparameter (Status Quo der Versorgungsqualität, Präferenzen von Versicherten und Patienten), können nur Daten aus dem Therapiealltag („real world data“ – RWD) liefern. Fehlende Ressourcen, ungelöste Probleme der Zugänglichkeit dieser Daten und eine teilweise unzureichend entwickelte oder verbreitete Methodik der Aufbereitung und Analyse dieser Daten tragen wesentlich zur suboptimalen Nutzung von RWD bei.

In einigen Ländern fokussiert HTA nicht auf die Bewertung einer spezifischen Technologie im Kontext eines spezifischen (nationalen) Gesundheitssystems, sondern ist im Wesentlichen mit der Neubewertung eher kontextunspezifischer Aspekte wie der relativen klinischen Wirksamkeit und Sicherheit betraut. Angesichts der Begrenztheit nationaler HTA-Ressourcen deutet die gleichzeitig in vielen EU-Mitgliedsstaaten durchgeführte Neubewertung klinischer Outcome-Parameter von Arzneimitteln auf der Basis der gleichen zum Zeitpunkt der Zulassung vorliegenden klinischen Studien auf ein beachtliches Potential für die internationale Zusammenarbeit in diesem Feld hin. Der grenzüberschreitenden Zusammenarbeit im Bereich der Evaluation der relativen klinischen Wirksamkeit von Gesundheitstechnologien hat sich das europäische HTA-Netzwerk EUnetHTA verschrieben [8], [9]. Mit den auf diese Weise eingesparten Ressourcen könnte nationales HTA sich stärker auf die Untersuchung und Optimierung der nationalen Versorgungsrealität konzentrieren, zu einer kontinuierlichen Verbesserung der Ergebnisqualität im Gesundheitswesen beitragen und so dem ursprünglichen Ideal von HTA wieder näherkommen.

Zunehmende Bedeutung von RWD in regulatorischen Zulassungsprozessen als Herausforderung für Preisfindungs- und Erstattungsentscheidungen

In den meisten Ländern wird das Potential von HTA als Grundlage von Desinvestitionsentscheidungen bei weitem nicht ausgeschöpft. Überwiegend wird HTA zur Unterstützung von Entscheidungen eingesetzt, die die Aufnahme neuartiger Leistungen, z.B. vor allem von Arzneimitteln, in den Leistungskatalog solidarisch finanzierter Gesund-

heitssysteme regulieren. Nur selten wird HTA zur Unterstützung von Optimierungsentscheidungen über den Einsatz vermutlich (partiell) obsoleter Leistungen eingesetzt. Um den Einsatz von Gesundheitsleistungen im Laufe der Zeit zu optimieren, muss sich HTA wesentlich stärker als heute auf die Erhebung und Nutzung von Daten aus dem realen Versorgungsgeschehen konzentrieren. Nur auf der Basis dieser Daten wird es möglich sein, die am Versorgungsgeschehen beteiligten Akteure, aktiv in einen kontinuierlichen an Behandlungsergebnissen orientierten Lernprozess einzubeziehen. In vielen Ländern werden mittlerweile Daten über die mit dem Einsatz von Arzneimitteln erzielten Behandlungsergebnisse erhoben. Im Zentrum der Nutzung dieser Daten steht häufig jedoch die Frage nach finanzieller Kompensation für nicht eingetretene Behandlungsergebnisse und nicht die von allen Beteiligten gemeinsam vorangetriebene Verbesserung der Versorgungsqualität. So wird häufig übersehen, dass informierte und motivierte Leistungserbringer und Patienten einen ganz wesentlichen Beitrag zur Erzielung nachhaltiger Behandlungserfolge leisten können und auch müssen.

Zunehmende Bedeutung von RWD in regulatorische Zulassungsprozessen als Herausforderung für Preisfindungs- und Erstattungsentscheidungen

Auch im Rahmen der von der europäischen Zulassungsbehörde EMA und den beteiligten nationalen Zulassungsbehörden angestrebten schrittweisen regulatorischen Zulassung für bestimmte Arzneimittel („adaptive licensing“ oder „adaptive pathways“) wird der prospektiv geplante Einsatz von Daten aus dem Therapiealltag zur Schließung bestehender Evidenzlücken nach initialer Zulassung eine zentrale Rolle für nachgelagerte Zulassungs- und Erstattungsentscheidungen spielen [10]. Wesentlich für den Erfolg dieses neuartigen, prospektiv geplanten schrittweisen Zulassungsprozesses wird die Entwicklung korrespondierender nationaler Preisfindungs- und Erstattungsprozesse sein. Nur wenn diese Prozesse die angestrebte frühere Verfügbarkeit bestimmter Arzneimittel für priorisierte Patientengruppen mit besonderem Therapiebedarf in schwer behandelbaren Indikationen adäquat reflektieren, wird sich das Konzept der schrittweisen regulatorischen Zulassung durchsetzen können. Der klinisch-regulatorische Kontext innerhalb dessen bestimmte Patientengruppen mit von gesellschaftlich legitimierten Entscheidungsträgern determiniertem hohem Therapiebedarf im Rahmen von prospektiv geplanten schrittweisen Zulassungsverfahren priorisiert werden, kann die Verfügbarkeit von klinischer Evidenz zum Zeitpunkt der initialen Zulassung oder auch grundsätzlich so einschränken, dass die den nachgelagerten Preisverhandlungen zugrundeliegenden hypothetischen maximalen Evidenzstandards nicht mehr erreicht werden können. Einen wesentlichen Beitrag könnte in diesem Zusammenhang der Übergang von hypothetisch maximalen Evidenzstandard zum Konzept der „kontext-spezifisch erwartbaren Evidenz“ im Rahmen von Preisfindungsprozessen

liefern. Gemäß diesem Konzept wird im Rahmen der Preisfindung nicht ein hypothetisch maximaler Evidenzstandard zugrunde gelegt, sondern ein „kontext-spezifisches“ Evidenzniveau, das angesichts des klinisch-regulatorischen Kontextes vom forschenden Pharma-Unternehmen zum Zeitpunkt der initialen Zulassung erwartet werden kann. Das Pharma-Unternehmen kann im Rahmen dieses Konzeptes zur weiteren Beibringung von Evidenz nach der initialen Zulassung verpflichtet werden, die je nach Fragestellung und praktischer Realisierbarkeit in weiteren randomisiert-kontrollierten Studien oder auch im Therapiealltag unter Routinebedingungen erhoben wird.

Für die erfolgreiche Einführung dieser neuartigen Zulassungsprozesse ist die konstruktive Zusammenarbeit aller beteiligten Stakeholder erforderlich. Nur so wird es möglich sein, in einem planvollen Prozess Konsequenzen für die beteiligten Parteien frühzeitig abzuschätzen und systematisch in einem umfassenden Evidenzentwicklungsplan zu adressieren. Dieser Plan beschreibt wie konventionelle klinische Studien, RWD und andere Verfahren der Evidenzgenerierung entscheidungsrelevante Fragen der beteiligten Entscheidungsträger (Zulassungsbehörde, HTA/Kostenträger, Leistungserbringer und Patienten) im Laufe des Produktlebenszyklus adressieren werden.

RWD als Grundlage bedingter Preisfindungs- und Erstattungsmodelle

In vielen Ländern werden innovative Vergütungsmodelle auf der Basis von RWD eingesetzt oder erprobt. Je nach Ausgestaltung erfolgt die Vergütung abhängig vom konkreten Einsatz eines Arzneimittels in bestimmten Patientengruppen, für die ein unterschiedlicher Mehrnutzen festgestellt wurde, oder aber auch abhängig von definierten Behandlungsergebnissen. Mit derartigen Vergütungsmodellen wird versucht, das Potential eines Arzneimittels, einen relativen Mehrnutzen im Sinne relevanter Behandlungsergebnisse zu generieren, möglichst differenziert zu honorieren. Eine so differenzierte Honorierung von Gesundheitsleistungen kann wesentlich zum ökonomisch optimalen Einsatz von Gesundheitsleistungen und -technologien beitragen.

HTA und Real World Data – Relevanz für den Gemeinsamen Bundesausschuss

- Matthias Perleth

Vorbemerkung

Bei „real world“- oder „real life“-Daten handelt es sich um solche Daten, die außerhalb der kontrollierten Bedingungen von prospektiven klinischen Studien generiert werden. Damit sollen Daten aus in der Regel randomisierten kontrollierten Studien (RCTs) ergänzt werden, da diese an selektierten homogenen Patientenpopulationen

unter Idealbedingungen durchgeführt werden, und ihnen somit eine fehlende Verankerung in der „echten“ Welt unterstellt wird. Dem kommt entgegen, dass zunehmend administrative Datenbestände (Sekundärdaten) erschlossen werden, aber auch Register (krankheits- oder prozedurspezifisch) eingerichtet werden, die ein Abbild der Versorgungsrealität enthalten. Drittens werden „real world“-Daten für Modellierungen verwendet, etwa im Kontext gesundheitsökonomischer Fragestellungen. Die International Society for Pharmacoeconomics and Outcomes Research (ISPOR) hat sogar eine Task Force gegründet, die 2007 eine ausführliche Publikation zu dem Thema vorlegte, wobei der Fokus auf Kostenübernahme- und Finanzierungsentscheidungen gerichtet wurde [11]. Diese führt an, „real world“-Daten würden besser Auskunft geben zur „effectiveness“ von medizinischen Methoden, der Wirksamkeit unter so genannten Alltagsbedingungen bzw. der Generalisierbarkeit von Studienergebnissen oder in der Formulierung der Task Force: „Yet, efficacy evidence in a particular group or subgroup is typically insufficient to project the size of the effectiveness impact in the population that would actually use a product.“

In diesem Beitrag geht es nicht darum, eine neue Definition zu präsentieren oder vorhandene kritisch zu diskutieren, sondern darzustellen, ob und in welchem Zusammenhang „real world“-Daten im Kontext der Entscheidungsfindung im Gemeinsamen Bundesausschuss (G-BA) eine Rolle spielen (können).

Hintergrund: Der Gemeinsame Bundesausschuss

Der Gemeinsame Bundesausschuss (G-BA) wurde 2004 als juristische Person des öffentlichen Rechts errichtet und konkretisiert den Leistungskatalog der GKV durch Entscheidungen zur Anwendbarkeit einzelner Technologien in Form von verbindlichen Richtlinien. Er hat weitgehende Zuständigkeit für die Bewertung von Arzneimitteln, nichtmedikamentösen Untersuchungs- und Behandlungsmethoden, die Entwicklung von Disease Management-Programmen, Qualitätssicherung der stationären und ambulanten Versorgung, der Psychotherapie, der Versorgung mit Heil- und Hilfsmitteln, Zahnmedizin und der Bedarfsplanung im ambulanten Sektor. Er wird durch Vertreter des GKV-Spitzenverbands, der Leistungserbringer (ambulant tätige Ärztinnen und Ärzte, Psychotherapeutinnen und Psychotherapeuten, Krankenhäuser und Zahnärztinnen und Zahnärzte) sowie durch Patientenvertreterinnen und -vertreter (mit beratender Stimme) gebildet (<https://www.g-ba.de/>).

Insbesondere für die Bewertung des Zusatznutzens von Arzneimitteln wird immer wieder gefordert, nicht nur Daten aus RCTs zu berücksichtigen. Seit 2011 hat der G-BA die gesetzliche Aufgabe, für alle neu zugelassenen Arzneimittel mit neuen Wirkstoffen sofort nach Markteintritt eine (Zusatz-)Nutzenbewertung durchzuführen (§ 35a SGB V). Deren Ergebnis ist die Entscheidungsgrundlage dafür, wieviel die gesetzliche Krankenversicherung nach entsprechenden Preisverhandlungen für ein neues Arz-

Tabelle 1: Evidenzquellen

Arzneimittel	RCTs, sonstige klinische Studien, Statistiken, epidemiologische Daten
Qualitätssicherung	Abrechnungsdaten, im Routinebetrieb erhobene Daten (Qualitätsindikatoren)
nicht-medikamentöse Untersuchungs- und Behandlungsmethoden	RCTs, sonstige klinische Studien, eventuell Registerdaten
Prävention, Impfungen	RCTs, sonstige klinische Studien, eventuell Modellierungen, epidemiologische Daten
Psychotherapie, Zahnmedizin	RCTs, sonstige klinische Studien
veranlasste Leistungen	RCTs, sonstige klinische Studien, Surveys, epidemiologische Studien
sektorenübergreifende Versorgung, Disease Management-Programme	RCTs, sonstige klinische Studien, Sekundärdaten
Bedarfsplanung	Statistiken, versorgungsepidemiologische Daten, ev. Modellierungen
übergreifend: Evaluationsprojekte	Surveys, Statistiken, Sekundärdaten, epidemiologische Daten, Krebsregister

neimittel mit einem neuen Wirkstoff zahlt. Im Grunde überprüft der G-BA den von den Herstellern behaupteten zusätzlichen Nutzen des neuen Arzneimittels im Vergleich zum bisherigen Standard, den diese im Rahmen eines eigens erstellten Dossiers darlegen und mit Studiendaten untermauern. Entscheidend ist dabei, ob das neue Medikament mit dem Behandlungsstandard (die zweckmäßige Vergleichstherapie) valide verglichen wurde. Hierfür sind RCTs der Standard. Dies und weitere Details sind in der Verfahrensordnung (VerfO) des G-BA geregelt [12].

Auch für die Bewertung von Untersuchungs- und Behandlungsmethoden sieht die VerfO vor, dass deren Nutzen in der Regel im Rahmen von RCTs nachgewiesen wird (2. Kapitel §13 Absatz 2). Für die Frage der Bewertung des patientenrelevanten Nutzens spielen „real world“-Daten demnach keine Rolle. Für die Entscheidungsfindung stellt sich zunächst die Frage nach der Generalisierbarkeit der Ergebnisse über die relevanten Studienpopulationen hinaus nicht. Fehlende Evidenz wird in der Regel konstatiert, wenn die Studienlage zum Nutzen einer Methode (oder eines Arzneimittels) unzureichend ist, d.h., es fehlen RCTs. Dies kann *nicht* durch administrative Daten, Registerdaten, Beobachtungsstudien, Surveys oder sonstige Daten ersetzt werden. Die Anzahl der so genannten „Aussetzungsbeschlüsse“, also der Beschlüsse, die ein Beratungsverfahren wegen fehlender Evidenz für einen bestimmten Zeitraum aussetzen, ist dementsprechend lang.

Dennoch stellt sich die Frage, ob die Ergebnisse aus RCTs auf die mutmaßliche Zielpopulation anwendbar sind. Windeler [13] hat hierzu unter dem Stichwort externe Validität das Nötige festgehalten: Die Zielpopulation (die „real world“) stellt die Grundgesamtheit dar, aus der die Studienpopulation rekrutiert wurde. Sofern keine gewichtigen Gründe dagegen sprechen, gilt die (inhaltliche, keine

statistische) Annahme, dass die Studienergebnisse auch für die Grundgesamtheit gelten, d.h. der Therapieeffekt erhalten bleibt. Die jeweils geltenden wesentlichen Bedingungen, unter denen die Therapieeffekte nachgewiesen wurden (und das ist bei nichtmedikamentösen Methoden relevanter als bei Arzneimitteln), werden ggf. als Bestandteil der Richtlinie festgeschrieben, um die Übertragbarkeit zu sichern (bspw. Strukturvorgaben und Qualitätsanforderungen). So genannte pragmatische RCTs mit sehr breit angelegten Einschlusskriterien bieten hinsichtlich der Frage der Praxistauglichkeit Vorteile und stellen somit einen möglichen Kompromiss dar.

Stellenwert von „Real world“-Daten im G-BA

Tabelle 1 stellt für die jeweiligen Zuständigkeitsbereiche dar, welche Evidenzquellen (neben RCTs) bei Beratungen und der Entscheidungsfindung eine Rolle spielen können. Dabei wird deutlich, dass eine Vielzahl von Datenquellen ausgewertet wird, von denen einige dem Bereich „real world“ zugeordnet werden können, bspw. Daten aus Krebsregistern oder epidemiologischen Studien. Diese Daten werden allerdings in der Regel abseits von Fragestellungen der Nutzenbewertung herangezogen, etwa im Kontext der Qualitätssicherung. Insbesondere im Rahmen von Evaluationsprojekten, die dazu dienen, die tatsächliche Umsetzung von Richtlinien des G-BA in der Praxis zu evaluieren und Änderungsbedarf zu identifizieren, werden „real world“-Daten benötigt.

Fazit

Im Rahmen der Nutzenbewertung von Arzneimitteln und nichtmedikamentösen Untersuchungs- und Behandlungsmethoden spielen „real world“-Daten in der Regel keine

Rolle. Fragestellungen jenseits der Nutzenbewertung werden allerdings mit Daten aus Alltagsbedingungen informiert. Aktuell ist nicht abzusehen, dass sich diese Situation grundsätzlich ändert. Lediglich so genannte pragmatische RCTs würden beide Welten im Rahmen der Nutzenbewertung annähern.

Soweit der Überblick zu HTA in Europa und dem Umgang mit Real World Data. Im Folgenden werden neue Aspekte der Verwendung von Surrogatparametern diskutiert, wobei insbesondere die Frage methodischer „Gaps“ aufgeworfen wird.

Surrogatvalidierung im HTA – Mind the GAP

- **Kirsten H. Herrmann**

Health Technology Assessments (HTA) gehören international und auch im deutschen Gesundheitswesen zu den Standardinstrumenten der Entscheidungsfindung insbesondere zur Definition des Leistungskatalogs der Krankenkassen. Stakeholder wie zum Beispiel Krankenkassen, Leistungserbringer und Patienten sind aus unterschiedlichen Perspektiven interessiert, Folgen von Gesundheitstechnologien abzuschätzen. Insbesondere epidemiologische und biometrische Fragen und Methoden spielen eine Rolle, daneben werden systematische Reviews, Nutzen-Schaden-Abwägungen, gesundheitsökonomische Evaluationen und entscheidungsanalytische Modellierungen eingesetzt [14], [1]. In unterschiedlichen Ländern werden diese Rollen unterschiedlich gewichtet. Vor diesem Hintergrund entwickeln sich zunehmende Bestrebungen nach Harmonisierung in der europäischen Zusammenarbeit. Seit vielen Jahren ist das „European network for Health Technology Assessment“ (EUnetHTA) im Bereich der Harmonisierung sehr aktiv [15]. Im Rahmen der Zusammenarbeit werden verschiedene Punkte kontrovers diskutiert. Die Entscheidungsmechanismen verschiedener Gesundheitssysteme (steuerfinanziert, marktwirtschaftlich oder versicherungsbasiert) unterliegen verschiedenen rechtlichen Rahmenbedingungen und folglich werden Möglichkeiten der Regulation, der Diffusion und Nutzung von Technologien unterschiedlich eingesetzt.

Von der Prioritätensetzung bei der Auswahl zu bewertender Technologien bis zur Methodik zur Erstellung von HTAs werden Schwerpunkte unterschiedlich gewichtet. Themen wie der methodologische Umgang mit Real World Data und Surrogatparametern werden seit vielen Jahren wiederkehrend diskutiert. Sichtweisen verschiedener Stakeholder können variieren und folglich kontrovers diskutiert werden. Eine frühe Studie folgert in Bezug auf klinische Studien, dass Surrogate einen vorläufigen Hinweis bieten können, der dann durch weitere Studien belegt werden muss, bevor Innovationen eingesetzt werden [16]. Ein neuerer Review untersucht den Nutzen von Surrogat Outcomes in Kosten-Effektivitäts-Modellen in HTAs in Großbritannien. In dieser Arbeit werden 35 Studi-

en eingeschlossen, von denen 4 Studien das Kosten-Effektivitäts-Modell auf Surrogat Outcomes basieren. Die Transparenz der Quantifizierung und Extrapolation von Unsicherheit von Surrogat vs. endgültiges Outcome variierte erheblich in diesen 4 Studien [17].

Am Beispiel der Surrogatparameter, die in den HTA-Verfahren vielfach im Bereich der Onkologie aktuell diskutiert werden, werden verschiedene Lösungsansätze durchgespielt. So werden Ansätze untersucht, um Regressionsverfahren als Ersatz für bivariate Modelle zur Bestimmung des Surrogatschwellenwerts bei korrelationsbasierten Validierungsverfahren einzusetzen. Die Autoren folgern, dass vereinfachte Modellansätze mit aggregierten Daten keine dem ursprünglichen bivariaten Modell auf Basis von individuelle Patientendaten gegenüber gleichwertige Ergebnisse liefern. Surrogate Threshold Effects, die sich unter Verwendung von Meta-Regressionen auf Basis aggregierter Daten ergeben, können jedoch in den betrachteten Fällen als konservativer Ersatz angesehen werden [18]. Ergebnisse von Simulationsstudien diskutieren die Möglichkeit der Surrogatvalidierung durch Korrelation und Surrogate Threshold Effect. Die durchgeführten Simulationen zeigen, dass die im Rapid Report beschriebene Methodik, wonach die untere Grenze des Konfidenzintervalls ausschlaggebend für eine hohe Korrelation bei der Surrogatvalidierung sein soll, eine in der Praxis kaum zu überwindende Hürde darstellt [19].

In einem weiteren Vortrag untersuchten Dr. Christoph Schürmann und Dr. Wiebke Sieben aus dem IQWiG, ob und welche Regressionsmodelle (einfache und gewichtete lineare Regression, Meta-Regression mit zufälligen Effekten) einen Ersatz darstellen für bivariate Modelle zur Bestimmung des Surrogatschwellenwerts bei korrelationsbasierten Validierungsverfahren.

Surrogatvalidierung durch Korrelation und Surrogate Threshold Effect – Ergebnisse von Simulationsstudien

- **Johanna Gillhaus, Ralf Goertz, Ulli Jeratsch, Friedhelm Leverkus**

Hintergrund

In onkologischen Studien wird oftmals statt des patientenrelevanten Endpunkts Gesamtüberleben (overall survival, OS) der Endpunkt progressionsfreies Überleben (progression-free survival, PFS) erfasst. Für eine Anerkennung von PFS als patientenrelevant im Verfahren der Nutzenbewertung gilt es, dieses als Surrogatendpunkt für OS in der betrachteten Indikation zu validieren. Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) hat im Rahmen eines Rapid Report Methoden zur Validierung von Surrogatendpunkten dargestellt und Empfehlungen zur Verwendung von korrelationsba-

sierten Verfahren ausgesprochen. In diesen Methoden werden die Einschätzung der Aussagesicherheit der Studienergebnisse und der Zusammenhang zwischen den Effektschätzern des Surrogats und des patientenrelevanten Endpunkts auf Studienebene einbezogen. Der Zusammenhang wird mit dem Korrelationskoeffizienten r bzw. dem Bestimmtheitsmaß R^2 und entsprechendem Konfidenzintervall (KI) gemessen. Für den Nachweis der Validität des Surrogats müssen eine hohe Korrelation sowie eine hohe Aussagesicherheit der Studienergebnisse vorliegen. Im Falle einer mittleren Korrelation kann das Konzept des Surrogate Threshold Effects (STE) zur Festlegung von Schwellenwerten für den Effektschätzer des Surrogatendpunkts angewandt werden.

Methoden

In Simulationsstudien wird nun untersucht, welche Bedingungen für eine erfolgreiche Surrogatvalidierung mit korrelationsbasierten Verfahren erfüllt sein müssen. Variierende Parameter sind die Effektschätzer des Surrogats und des patientenrelevanten Endpunkts, die Korrelation zwischen den Effektschätzern, die Patientenzahl sowie die Anzahl der Studien. Es wird analysiert, in welchen Szenarien der Nachweis einer hohen Korrelation gelingt und falls nicht, welche Voraussetzungen vorliegen müssen, dass anhand des Surrogats unter Einbeziehen des STE-Konzepts noch Schlüsse auf den patientenrelevanten Endpunkt gezogen werden können. Die Herausforderungen der vom IQWiG präferierten Methodik zur Surrogatvalidierung in der Praxis werden analysiert.

Ergebnisse

Die Simulation der Surrogatvalidierung über das korrelationsbasierte Verfahren sowie die analytische Herleitung der Power zeigen, dass diese bei moderater Studienanzahl und starker zugrundeliegender wahrer Korrelation dennoch sehr gering ist. Die Power liegt für $n=5$ Studien und Korrelation $\rho=0,9$ unter 6%. Es wäre eine sehr hohe Korrelation von $\rho=0,95$ in mindestens $n=25$ Studien erforderlich, um eine Power von 80% zu erhalten. Dieses Szenario ist in der Realität allerdings völlig unplausibel. In der Simulation zur Anwendung des STE-Konzepts liegt die Power nur bei etwa ein Fünftel der betrachteten Szenarien über 80%. Dabei handelte es sich jedoch um Szenarien, in denen die Parameterkonstellationen aus hoher Studien- und Patientenzahl und großem Effekt des OS eher unrealistisch sind. Der Parameter der Korrelation ρ zwischen den Effektschätzern der Studien wirkt sich ebenso wie verschiedene Effekte des PFS kaum auf die Power des STE-Verfahrens aus.

Fazit

Die durchgeführten Simulationen zeigen, dass die im Rapid Report beschriebene Methodik, wonach die untere Grenze des Konfidenzintervalls ausschlaggebend für eine hohe Korrelation bei der Surrogatvalidierung sein soll,

eine in der Praxis kaum zu überwindende Hürde darstellt. Bei gering bis moderat angenommener Studienanzahl – wie es für eine Validierung von Surrogatendpunkten im Rahmen der frühen Nutzenbewertung realistisch erscheint – ist die Power selbst bei hoher, wahrer Korrelation äußerst gering. Problematisch erscheint weiterhin die Empfehlung, die Aussagekraft der Studien in die Analyse mit einzubeziehen, auch wenn dies prinzipiell gerechtfertigt erscheint. Bei Betrachtung der Definition des Korrelationskoeffizienten und dessen Dichtefunktion wird zudem klar, dass die empirische Korrelation unter Annahme einer festen wahren Korrelation gar nicht von der Varianz der Einzelschätzer, sondern nur von der Anzahl der Wertepaare abhängt. Die Patientenzahl hat somit keine Auswirkung auf das Konfidenzintervall der Korrelation. Dies gilt ebenso, wenn Modelle mit Gewichtung der Studien verwendet werden. Die Anwendung des STE-Konzeptes gemäß der im Rapid Report beschriebene Methodik ist ebenfalls zu restriktiv. Ein Vergleich des STE mit dem Punktschätzer des Surrogatendpunkts wäre eine Alternative, die in realistischen Szenarien sehr geringe α -Fehler zeigte. Gegebenenfalls kann in weiteren Simulationen überprüft werden, ob ein vorgeschalteter Test (z. B. ob der ein- σ -Bereich der Verteilung des $HR_{0,5}$ unter 1 liegt) den α -Fehler auch in diesen unrealistischen Szenarien kontrolliert.

Regressionsverfahren als Ersatz für bivariate Modelle zur Bestimmung des Surrogate Threshold Effects (STE) bei korrelationsbasierten Validierungsverfahren

- Christoph Schürmann, Wiebke Sieben

Der Einsatz von Surrogatendpunkten in klinischen Studien mag vielversprechend sein, wenn beispielsweise die lange Zeit bis zum Eintreten des eigentlich interessierenden Ereignisses durch die Beobachtung eines zeitig früher eintretenden Ereignisses verkürzt werden kann. Insbesondere in der Onkologie werden mit diesem Ziel Surrogatendpunkte wie das progressions- oder krankheitsfreie Überleben erhoben, um damit Rückschlüsse auf das eigentlich interessierende Gesamtüberleben im Vergleich zweier Interventionen zu erlauben. Dies erfordert jedoch eine methodisch angemessene Validierung der verwendeten Surrogatendpunkte. Unter den hierfür vorhandenen Ansätzen sind insbesondere korrelationsbasierte Verfahren ([20], [21] und darauf aufbauende Weiterentwicklungen) zu nennen, die in einem bivariaten statistischen Modell den Zusammenhang von Surrogat- und eigentlichem Endpunkt auf individueller und Studienebene analysieren. Hierzu sind allerdings individuelle Patientendaten (IPD) notwendig. Als Validierungsmaße können aus diesen Modellen einerseits Korrelationsmaße wie der Korrelationskoeffizient (auf Patienten- und Studienebene) oder das Bestimmtheitsmaß abgeleitet werden, anderer-

seits auch Schwellenwerte für Effekte auf den Surrogatendpunkt, aus deren Über- oder Unterschreiten in einer neuen klinischen Studie sich ein Effekt auf den eigentlichen Endpunkt ableiten lässt. Derartige Surrogate Threshold Effects (STEs) bieten in der Anwendungssituation den Vorteil, die Stärke des Behandlungseffekts für den eigentlichen Endpunkt besser charakterisieren zu können, als dies durch Korrelationsmaße möglich ist. Die Klasse dieser bivariaten Modelle ist in der Literatur mittlerweile intensiv untersucht und für einige Indikationen und Interventionen auch beispielhaft angewendet worden [22], [23], [24]. Daneben finden sich Anwendungen, die in Ermangelung individueller Patientendaten auf Regressionsmodelle auf Basis aggregierter Daten zur Beurteilung der Surrogatvalidität zurückgreifen [25], [26], [27]. Bei diesen Ansätzen ist zunächst fraglich, ob diese zu gleichen Resultaten hinsichtlich des STEs wie die bivariaten Modelle auf IPD-Basis führen.

In einer Simulationsstudie [28] wurden individuelle Patientendaten für zwei stetige Endpunkte generiert, wobei die Zahl der zur Validierung eingesetzten Studien, die Studienfallzahl, die Größe der Behandlungseffekte sowohl auf Surrogat- und eigentlichen Endpunkt und die Korrelationen der Endpunkte auf individueller und Studienebene systematisch variiert wurden. Für jeden so erzeugten Datensatz wurden anschließend die aus dem ursprünglichen, auf IPD basierenden bivariaten Modell ermittelten STEs und solche verglichen, die sich aus vereinfachten Modellen wie linearem Regressionsmodell, linearem gewichtetem Regressionsmodell und Meta-Regressionsmodell mit zufälligen Effekten auf Basis von aggregierten Daten ergeben. Unabhängig von den genannten Einflussfaktoren zeigt sich dabei, dass die linearen Regressionsmodelle nahezu ausschließlich zu STEs führen, die betragsmäßig kleiner als die des ursprünglichen bivariaten Modellansatzes mit IPD sind, und diese damit unterschätzen. Umgekehrt liegen die STEs nach Meta-Regression fast immer oberhalb der Vergleichswerte. Die Stärke des Unterschieds wird sowohl absolut als auch relativ im Wesentlichen durch die Korrelation der Behandlungseffekte auf Studienebene bestimmt, daneben auch durch die Anzahl der Studien des Validierungsdatensatzes. Die übrigen untersuchten Faktoren hatten in den Modellen diesbezüglich keinen relevanten Einfluss. Daraus lässt sich folgern, dass vereinfachte Modellansätze mit aggregierten Daten keine dem ursprünglichen bivariaten Modell auf Basis von IPD gegenüber gleichwertige Ergebnisse liefern. STEs, die sich unter Verwendung von Meta-Regressionen auf Basis aggregierter Daten ergeben, können jedoch in den betrachteten Fällen als konservativer Ersatz angesehen werden.

Anmerkung

Interessenkonflikte

Die Autoren erklären, dass keine Interessenkonflikte in Zusammenhang mit diesem Artikel vorliegen.

Literatur

1. Perleth M, Busse R, Gerhardus A, Gibis B, Lühmann D, Zentner A, editors. Health Technology Assessment: Konzepte, Methoden, Praxis für Wissenschaft und Entscheidungsfindung. 2nd ed. Berlin: Med. Wiss. Verl.-Ges; 2014.
2. Shemilt I, Mugford M, Vale L, Marsh K, Donaldson C. Evidence-based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice, 2nd edition. Chichester: Wiley-Blackwell; 2010. DOI: 10.1002/9781444320398
3. Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. Official Journal of the European Union. 2011 [cited 2016 Jun 16]; L88/45:45-65. Available from: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:088:0045:0065:EN:PDF>
4. European Commission. Health Technology Assessment Network. [cited 2016 Jun 16]. Available from: http://ec.europa.eu/health/technology_assessment/policy/network/index_en.htm
5. EU Health Technology Assessment Network. Strategy for EU Cooperation on Health Technology Assessment. 2014 [cited 2016 Jun 16]. Available from: http://ec.europa.eu/health/technology_assessment/docs/2014_strategy_eucooperation_hta_en.pdf
6. European Commission. HTA Network reflection paper on "reuse of joint work in national HTA activities". 2015 [cited 2016 Jun 16]. Available from: http://ec.europa.eu/health/technology_assessment/docs/reuse_jointwork_national_hta_activities_en.pdf
7. EUnetHTA. EUnetHTA Guidelines. 2016 [cited 2016 Jun 16]. Available from: <http://www.eunetha.eu/eunetha-guidelines>
8. Kleijnen S, George E, Goulden S, d'Andon A, Vitre P, Osinska B, Rdzany R, Thirstrup S, Corbacho B, Nagy BZ, Leufkens HG, de Boer A, Goettsch WG. Relative effectiveness assessment of pharmaceuticals: similarities and differences in 29 jurisdictions. Value Health. 2012 Sep-Oct;15(6):954-60. DOI: 10.1016/j.jval.2012.04.010
9. EUnetHTA. Joint Assessments. 2016 [cited 2016 May 3]. Available from: <http://www.eunetha.eu/joint-assessments>
10. Eichler HG, Baird LG, Barker R, Bloechl-Daum B, Børlum-Kristensen F, Brown J, Chua R, Del Signore S, Dugan U, Ferguson J, Garner S, Goettsch W, Haigh J, Honig P, Hoos A, Huckle P, Kondo T, Le Cam Y, Leufkens H, Lim R, Longson C, Lumpkin M, Maraganore J, O'Rourke B, Oye K, Pezalla E, Pignatti F, Raine J, Rasi G, Salmonson T, Samaha D, Schneeweiss S, Siviero PD, Skinner M, Teagarden JR, Tominaga T, Trusheim MR, Tunis S, Unger TF, Vamvakas S, Hirsch G. From adaptive licensing to adaptive pathways: delivering a flexible life-span approach to bring new drugs to patients. Clin Pharmacol Ther. 2015 Mar;97(3):234-46. DOI: 10.1002/cpt.59
11. Garrison LP Jr, Neumann PJ, Erickson P, Marshall D, Mullins CD. Using real-world data for coverage and payment decisions: the ISPOR Real-World Data Task Force report. Value Health. 2007 Sep-Oct;10(5):326-35. Available from: https://www.ispor.org/workpaper/RWD_TF/ISPORRealWorldDataTaskForceReport.pdf
12. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses in der Fassung vom 18. Dezember 2008. [zuletzt geändert am 18. Dezember 2014, veröffentlicht im Bundesanzeiger BAnz AT 15.04.2015 B2; cited 2016 Jun 16]. Available from: https://www.g-ba.de/downloads/62-492-1002/VerfO_2014-12-18_iK-2015-04-16.pdf
13. Windeler J. Externe Validität. Z Evid Fortbild Qual Gesundheitswesen. 2008 [cited 2016 Jun 16];102(4):253-9. Available from: <http://www.ebm-netzwerk.de/site/was-ist-ebm/images/windeler-externe-valid.pdf>

14. Cochrane Deutschland. Health Technology Assessment (HTA). 2016 [cited 2016 Feb 6]. Available from: <http://www.cochrane.de/de/hta>
15. EUnethTA. European network for Health Technology Assessment. 2016 [cited 2016 Feb 6]. Available from: <http://www.eunethta.eu/>
16. Gøtzsche PC, Liberati A, Torri V, Rossetti L. Beware of surrogate outcome measures. *Int J Technol Assess Health Care*. 1996;12(2):238-46. DOI: 10.1017/S0266462300009594
17. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. *Int J Technol Assess Health Care*. 2009 Jan;25(1):6-13. DOI: 10.1017/S0266462309090023
18. GMDS 2015: 60. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS); 06.09. - 09.09.2015; Krefeld. Düsseldorf: German Medical Science GMS Publishing House; 2015. Available from: <http://www.egms.de/de/meetings/gmgs2015/>
19. Buncke J, Goertz R, Jeratsch U, Leverkus F. Surrogatvalidierung durch Korrelation und Surrogate Threshold Effect: Ergebnisse von Simulationsstudien. GMDS 2015: 60. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS); 06.-09.09.2015; Krefeld. Düsseldorf: German Medical Science GMS Publishing House; 2015. DocAbstr. 139. DOI: 10.3205/15gmgs117
20. Burzykowski T, Molenberghs G, Buyse M, Geys H, Renard D. Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *J R Stat Soc Ser C Appl Stat*. 2001;50(4):405-22. DOI: 10.1111/1467-9876.00244
21. Burzykowski T, Molenberghs G, Buyse M, editors. The evaluation of surrogate endpoints. New York: Springer; 2005. (Statistics for Biology and Health). DOI: 10.1007/b138566
22. Burzykowski T, Molenberghs G, Buyse M. The validation of surrogate end points by using data from randomized clinical trials: A case-study in advanced colorectal cancer. *J R Stat Soc Ser A Stat Soc*. 2004;167(1):103-24. DOI: 10.1111/j.1467-985X.2004.00293.x
23. Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent DJ, Miller LL, Elfring GL, Pignon JP, Piedbois P. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol*. 2007 Nov;25(33):5218-24. DOI: 10.1200/JCO.2007.11.8836
24. Oba K, Paoletti X, Alberts S, Bang YJ, Benedetti J, Bleiberg H, Catalano P, Lordick F, Michiels S, Morita S, Ohashi Y, Pignon JP, Rougier P, Sasako M, Sakamoto J, Sargent D, Shitara K, Cutsem EV, Buyse M, Burzykowski T; GASTRIC group. Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: a meta-analysis. *J Natl Cancer Inst*. 2013 Nov;105(21):1600-7. DOI: 10.1093/jnci/djt270
25. Miksad RA, Zietemann V, Gothe R, Schwarzer R, Conrads-Frank A, Schnell-Inderst P, Stollenwerk B, Siebert U. Progression-free survival as a surrogate endpoint in advanced breast cancer. *Int J Technol Assess Health Care*. 2008;24(4):371-83. DOI: 10.1017/S0266462308080495
26. Flaherty KT, Hennis M, Lee SJ, Ascierto PA, Dummer R, Eggermont AM, Hauschild A, Kefford R, Kirkwood JM, Long GV, Lorigan P, Mackensen A, McArthur G, O'Day S, Patel PM, Robert C, Schadendorf D. Surrogate endpoints for overall survival in metastatic melanoma: a meta-analysis of randomised controlled trials. *Lancet Oncol*. 2014 Mar;15(3):297-304. DOI: 10.1016/S1470-2045(14)70007-5
27. Sidhu R, Rong A, Dahlberg S. Evaluation of progression-free survival as a surrogate endpoint for survival in chemotherapy and targeted agent metastatic colorectal cancer trials. *Clin Cancer Res*. 2013 Mar;19(5):969-76. DOI: 10.1158/1078-0432.CCR-12-2502
28. Schürmann C, Sieben W. Differences in surrogate threshold effect estimates between original and simplified correlation-based validation approaches. *Stat Med*. 2016 Mar;35(7):1049-62. DOI: 10.1002/sim.6778

Korrespondenzadressen:

Alric Rüther
IQWiG – Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Im Mediapark 8, 50670 Köln, Deutschland
alric.ruther@iqwig.de
Kirsten H. Herrmann
Bristol-Myers Squibb GmbH & Co. KGaA, Arnulfstr. 29, 80636 München, Deutschland
khtmlforme@gmail.com

Bitte zitieren als

Rüther A, Herrmann KH, Hebborn A, Perleth M, Schwarzer R, Schürmann C, Sieben W, Gillhaus J, Goertz R, Jeratsch U, Leverkus F, Schramm W. HTA und aktuelle Herausforderungen: Harmonisierung, Real World Data und Surrogatparameter. *GMS Med Inform Biom Epidemiol*. 2018;14(1):Doc02.
DOI: 10.3205/mibe000180, URN: urn:nbn:de:0183-mibe0001807

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/mibe/2018-14/mibe000180.shtml>

Veröffentlicht: 16.02.2018

Copyright

©2018 Rüther et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.