PD Dr. rer. pol. Markus Lüngen

Kosten-Effektivitäts-Analyse und Steuerung des Gesundheitswesens

Ein wesentlicher Forschungs- und Anwendungszweig der Gesundheitsökonomie ist die Durchführung von Kosten-Effektivitäts-Analysen. Diese werden oftmals als Entscheidungshilfe für die Bewertung neuer oder auch bestehender Therapien eingesetzt. Der nachfolgende Beitrag setzt sich mit den Folgen auseinander, welche die zunehmende Bedeutung von Kosten-Effektivitäts-Analysen für die Steuerung des Gesundheitswesens haben kann.

>> Die Kosten-Effektivitäts-Analyse (im folgenden CEA) hat in den vergangenen Jahrzehnten eine fulminante Karriere durchlebt, sowohl in der wissenschaftlichen Diskussion, als auch im Hinblick auf Qualität und Häufigkeit der Durchführung. Die Ursachen hierfür sind vielfältig. Wesentlich dürfte sein, dass verbesserte elektronische Datenverfügbarkeit und darauf aufbauende Auswertungsmöglichkeiten sich wechselseitig positiv beeinflusst haben. Dies wiederum ermöglichte praxistaugliche gesetzliche Vorgaben im Rahmen der Arzneimittelzulassung, was sowohl statistische Ansätze und Modellierungen, als auch das Angebot entsprechender Computerprogramme nochmals beflügelte.

Die Steuerung eines Gesundheitswesens über Kosten-Effektivitäts-Abwägungen spielt sich entgegen häufiger intuitiver Vermutungen nicht im wertfreien Raum ab. Es werden vielmehr fundamentale normative Festlegungen notwendig, um beispielsweise eine Größe wie 10.000 Euro pro zusätzlich gewonnenem Lebensjahr aus gesundheitspolitischer Sicht zu interpretieren.

Nachfolgend sollen die Grundzüge von Möglichkeiten und Grenzen der CEA im Hinblick auf die Steuerung von Gesundheitssystemen diskutiert werden. Insbesondere soll dabei auch die Einbeziehung von QALYs (qualitätsadjustierte Lebensjahre) diskutiert werden.

Ziel und Methode

Es wäre leicht, mehrere Seiten mit Problemen zu füllen, die Gesundheitsökonomen bei der praktischen Durchführung von CEA zu schaffen machen können. Diese Probleme reichen von fehlenden Daten über widersprüchliche Daten bis zu kreativen Änderungsvorschlägen für Auswertungspläne, nachdem die Daten gesichtet wurden. Die Geschichte der CEA könnte als eine Geschichte des „creative writing" erzählt werden. Auf diese Praxisprobleme soll hier nicht abgestellt werden. Das Deutsche Netzwerk Versorgungsforschung möchte hierzu in naher Zukunft praxisnahe Lösungen und Empfehlungen erarbeiten. Übersichten verwandter Themen existieren bereits, etwa für die Darstellung der Ergebnisse klinischer Studien (http://www.consort-statement.org/, Abfrage Nov 2008).

Ziel des Beitrags ist vielmehr, die Grundmuster der Wirkung und die zugrunde liegende ethisch-normativen Verankerung von CEA in der Gesundheitspolitik und -versorgung zu finden.

Aufgrund einer (nicht systematischen) Literatursuche wurden die Beiträge von Sculpher et al. (2005), Kind (2005), Gyrd-Hansen (2005) und Roberts et al. (2004) ausgewählt, da diese aus meiner Sicht einen ausgewogenen Überblick über den Stand der Problemlage bei CEA wieder geben. Einen guten Überblick über die deutsche Diskussion gibt das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) auf seiner Internetseite. (Methodenpapier zur Kosten-Nutzen-Bewertung http://www.iqwig.de/index.805.html Abfrage Nov 2008)

Ergebnisse

Es fiel zunächst auf, dass insbesondere in England und den USA die kritische Adaption des CEA-Ansatzes sehr viel verbreiteter ist als in Deutschland (siehe etwa Steinbrook (2008) für eine neuere zusammenfassende Sicht auf NICE). Während dort die Diskussion, häufig im Zusammenhang mit dem National Institute of Clinical Excellence (NICE), oftmals auch die gesellschaftliche Dimension aufgreift und einen ethischen Konsens anmahnt, gelangt die deutsche Diskussion, etwa um Methodenpapiere des IQWiG, kaum über ein Verhakeln um technische Details hinaus. Die eigentliche Frage, was CEA kann und was nicht, blieb häufig unberührt.

Eine zweite Auffälligkeit war, dass die grundlegende Kritik an der Eignung von CEA im Gesundheitswesen wohl genau so alt ist wie die Anwendung der CEA selbst. Dies kann als Beleg angesehen werden für eine lebhafte wissenschaftliche Diskussion, oder aber den Verdacht nähren, dass bereits zu Beginn der Nutzung von CEAs im Gesundheitswesen etwas grundlegend Kritikwürdiges nicht ausgeräumt werden konnte. Der Hinweis, dass CEA nun immerhin bereits über 30 Jahre durchgeführt werden, kann meines Erachtens in diesem Zusammenhang weder als Beleg für einen Erfolg noch als Beleg für eine Verirrung angesehen werden.

Abstract

Die Kosten-Effektivitäts-Analyse hat als Hilfe zur Entscheidungsfindung in den vergangenen Jahren eine enorme Aufwertung erfahren. Ihre auch gesetzliche Verankerung in der Wirtschaftlichkeitsprüfung führt dazu, dass die Belastbarkeit der Aussagen geprüft werden muss. Politische und methodische Skepsis richtet sich insbesondere gegen eine einfache Entscheidung anhand von monetären Kosten pro QALY einer Therapie. Stärker wird zudem die Diskussion, ob nicht subjektive Nutzeneinschätzungen des Einzelnen Vorrang haben sollen vor objektiven Einschätzungen einer Gruppe. Welche der Steuerungen im Gesundheitswesen endgültig gewählt werden, ist noch offen. Doch bedürfen beide Ansätze der methodischen Verfeinerung, bevor sie der Gesundheitspolitik hilfreiche Unterstützung anbieten können.

Schlüsselbegriffe

Kosten-Nutzen-Analyse, Wohlfahrtsökonomie, Gesundheitspolitik

Normative Verankerung in Utilitarismus und Liberalismus

Ein erster Problembereich in der Literatur behandelt die Frage, ob CEAs generell geeignet sind, eine Steuerung von Leistungen in der Gesundheitsversorgung angemessen vorzunehmen. Roberts et al. (2004) merken dazu an, dass sich aus der Ethik zwei Grundgebäude der Steuerung im Gesundheitswesen benennen lassen, in die CEAs unterschiedlich eingepasst werden können. (Siehe auch Abb. 1)

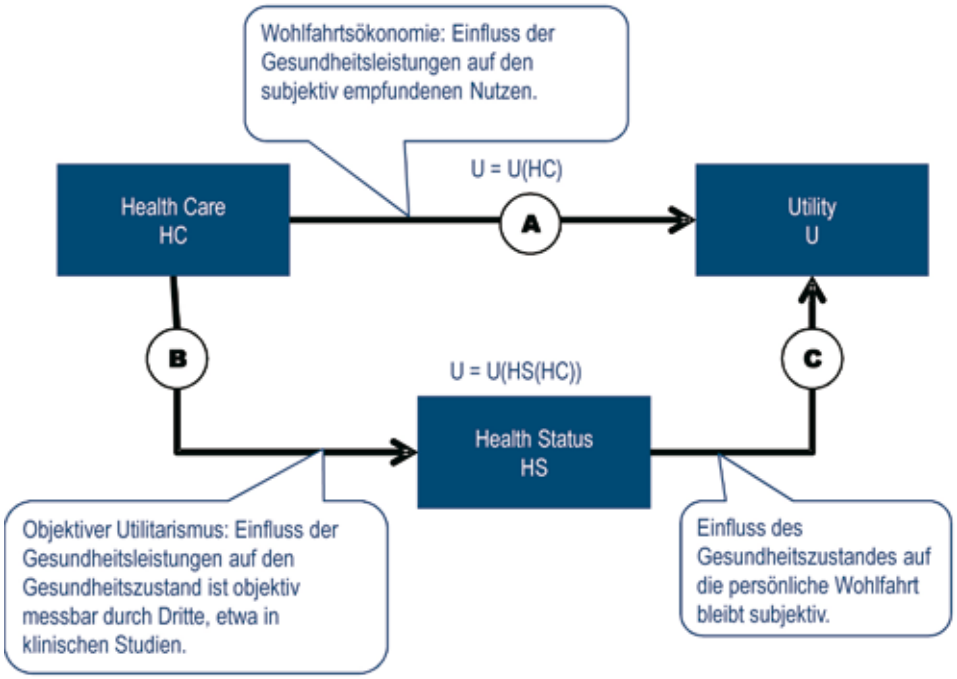


Abb.1: Zusammenhang zwischen Gesundheitsversorgung (Health Care), Nutzen (Utility) und Gesundheitszustand (Health Status) in der Wohlfahrtsökonomie und objektivem Utilitarismus.

Der **Utilitarismus** zielt auf die Konsequenzen des Handelns ab. Ziel ist bei ihm die Maximierung des Gesamtnutzens in der Gesellschaft („Das größte Glück der größten Zahl", siehe wesentlich hierzu wie immer Bentham, 1748-1832). Der **subjektive Utilitarismus** unterstellt, dass der einzelne Bürger am besten beurteilen kann, was ihm nutzt. Daher kann auch nur der einzelne Bürger über seine Handlungen zur Maximierung des Nutzens entscheiden. Unsicherheit über den Nutzen sowie eine Diskontierung wird beim subjektiven Utilitarismus unmittelbar innerhalb der geäußerten Zahlungsbereitschaft (Willingness-to-Pay; WtP) aus Sicht des Einzelnen berücksichtigt. Der bevorzugt vorgeschlagene Koordinationsmechanismus des Utilitarismus ist (aus praktischen Gründen) der Markt mit Wettbewerb. Die Analyse erfolgt im Rahmen der Wohlfahrtsökonomie. Weitgehend ungelöst bleibt innerhalb des subjektiven Utilitarismus das Problem der Gleichheit der Versorgung. Eine CEA kann innerhalb des subjektiven Utilitarismus dazu dienen, dem Patienten oder Nachfrager eine Entscheidungshilfe zu geben. Die Entscheidung über eine Inanspruchnahme selbst fällt er jedoch (basierend auf Präferenzen und Zahlungsbereitschaft) selbst. Ob der Patient eine CEA überhaupt begrüßt und sich an ihr in seiner Handlung ausrichtet, bleibt innerhalb des Gedankengebäudes des subjektiven Utilitarismus unbestimmt. Ebenso kann er sich von Werbung oder Gesprächen mit den Nachbarn beeinflussen lassen.

Der **objektive Utilitarismus** basiert ebenfalls auf der Annahme, dass der Nutzen maximiert werden soll, jedoch wird die Auswahl der nutzenstiftenden Maßnahmen durch eine Gruppe getroffen, und nicht vom Einzelnen. Diese Gruppe (etwa Regierung, Ärztekammer oder Krankenkassen) soll ihre Entscheidung bevorzugt auf der Grundlage von Evidenz und rationalen Zielen treffen. Aus dem objektiven Utilitarismus lässt sich beispielsweise die Forderung der Steuerung des Gesundheitswesens auf der Basis von QALYs ableiten. Sculpher et al. (2005) sehen die Abkehr von der Wohlfahrtsökonomie innerhalb des objektiven Utilitarismus durchaus positiv. Diese Abkehr bedeutet neben der Abkehr vom Subjektivismus auch eine Abkehr vom sogenannten Pareto-Kriterium für Änderungen des Status-Quo (und damit einer gewissen Lähmung der Debatte), eine Abkehr von den eventuell unerwünschten Einflüssen der bestehenden Einkommensverteilung auf die Inanspruchnahme, und schließlich eine weitgehende Abkehr von irrationalen Handlungen Einzelner auf zudem womöglich nicht funktionierenden Märkten.

Der **Liberalismus** schließlich argumentiert nicht mit den Zielen (also der Maximierung des Nutzens), sondern mit den Rechten des Einzelnen. Die Anhänger des Liberalismus denken insbesondere an das Recht, alles tun zu können, solange man andere nicht belästigt. Dies mündet häufig in der Forderung eines möglichst reduzierten Staates, der insbesondere auch nicht das Recht zur Verwendung des eigenen Einkommens beschneiden soll. Egalitärer Liberalismus wiederum argumentiert differenzierter, indem die Notwendigkeit von Ressourcen zur Wahrnehmung der eigenen Rechte betont werden. Daraus abgeleitet werden meist Umverteilung und die Vorhaltung von Grundausstattungen.

Aus diesen Grundmustern ließe sich ableiten, dass CEA entweder dem Einzelnen im Sinne des subjektiven Utilitarismus dienen kann, um die subjektive Nutzenmaximierung auf individueller Ebene rationaler zu gestalten. Oder aber CEA dient dazu, einer Gruppe den ihr zugewiesenen Auftrag zur objektiv verankerten Maximierung des Nutzen besser erfüllen zu können. Im Liberalismus schließlich können CEAs dazu dienen, Umfang und Kosten einer Grundversorgung abzuschätzen.

Je nach normativer Verankerung ergeben sich für Gegenstand, Perspektive und auch Einschlusskriterien der Patienten für CEAs erheblich abweichende Vorgaben. In der Regel (und insbesondere in Deutschland), überwiegt derzeit die Anwendung im Sinne des objektiven Utilitarismus, verbunden mit dem Ziel zur Formulierung einer Grundausstattung an Leistungen. Festzuhalten bleibt somit, dass CEA derzeit kaum im wohlfahrtsökonomischen Sinne eingesetzt wird.

Sollte diese Vereinigung von Wohlfahrtsökonomie und CEA zukünftig stärker betont werden (wofür wenig spricht), sind nicht unerhebliche methodische Hürden zu überwinden (siehe etwa den Überblick bei Gyrd-Hansen, 2005). Inwieweit so eine Zusammenführung überhaupt sinnvoll ist, bleibt zudem strittig. Es scheint eher eine Tendenz zu geben, CEAs auf ihren Ursprung in der Entscheidungstheorie zurück zu führen.

Unübersehbar sind jedoch Tendenzen, Gesundheitsversorgung für Patienten wettbewerblicher über Märkte zu steuern, also ein genereller Vormarsch der eher auf den subjektiven Nutzeneinschätzungen basieren-

den Steuerung. Welche Bedeutung CEA in diesem Umfeld haben können (etwa eine Art Stiftung Warentest für neue Therapien), ist ungewiss. Vorschläge wie eine Zusatzversicherung für innovative Therapien oder der massive Ausbau von Individuellen Gesundheitsleistungen (IGeL) lassen CEA nicht mal mehr zur Definition des relevanten Marktes (etwa der Kosten-wirksamen Therapien) zu. Vielmehr steht der Patient als Konsument im Vordergrund, welcher nach seinem subjektiven Nutzenempfinden auch eine objektiv nicht wirksame Leistungen nachfragen kann.

Wie auch immer die Zukunft der CEA aussehen wird, den politischen Prozess und die normative Diskussion über die gesellschaftlichen Ziele wird sie nicht ersetzen können. Gerade das Technische und scheinbar Wertfreie der CEA wird ihr immer stärker zum Verhängnis, je größer ihre formale Bedeutung im Gesetz wurde.

Herausforderungen bei Nutzenmessung in CEA

CEA bestehen grundsätzlich aus der Gegenüberstellung von Kosten- und Nutzendifferenzen zwischen Alternativen der Behandlung (beziehungsweise Nichtbehandlung). Dies wird ausgedrückt in der ICER (Incremental Cost Effectiveness Ratio), also den Mehrkosten pro zusätzlich gewonnener Nutzeneinheit. Auch wenn die Messung der Kosten eine Herausforderung darstellt, ist zumindest die Kosteneinheit unumstritten. Beim Nutzen verhält es sich anders. Die klassische Einteilung der CEA in ihre Untergruppen ist unmittelbar aus den verschiedenen Nutzengrößen abgeleitet. Insbesondere natürliche Größen (etwa Lebensjahre, Blutdrucksenkung) in der CEA im engeren Sinne und qualitätsadjustierte Lebensjahre (QALYs) in der Kosten-Nutzwert (cost-utility)-Analyse werden unterschieden.

Der Vorteil von QALYs liegt darin, dass sie zunächst nicht vergleichbar erscheinende Gesundheitsprogramme und Interventionen zu Allokationszwecken leichter handhabbar machen können. Ein QALY wird immer als ein QALY angesehen. Während der Nutzen einer Einheit Blutdrucksenkung kaum mit dem Nutzen einer vermiedenen Wiedereinweisung vergleichbar ist, können gewonnene QALYs über verschiedenste Gesundheitsprogramme verglichen werden. Die aus CEAs ableitbaren Ranglisten mit Angaben der Kosten pro gewonnenem QALY sollen dann dazu dienen, die Allokation eines Budgets rationaler zu gestalten (siehe Listen etwa bei Drummond, 1997).

Diese angenehme Perspektive bei der Nutzung von QALYs führte dazu, dass Gesundheitsökonomen einem erheblichen Druck ausgesetzt waren, diese Nutzengröße in ihrem Studiendesign als Endpunkt vorzusehen (Kind, 2005). QALYs als universelle Nutzeneinheit sind jedoch kein Nutzen im Sinne des subjektiven Utilitarismus, da bei QALYs nicht jeder Patient subjektiv den Nutzen einer Intervention bewertet, sondern eine (idealerweise repräsentative) Gruppe von Befragten stellvertretend für die Gesellschaft. Auch eine Aggregation von QALYs über Patientenfälle hinweg ist mit dem subjektiven Utilitarismus nicht vereinbar.

Neben dieser generellen Einschränkung stellen sich bei der Nutzung eines Nutzenindex und dessen Bewertung mit Kosten mehrere Fragen (siehe hierzu auch Roberts et al., 2004):

- Ist es gleich zu werten, ob 10 Jahre mit geringer Lebensqualität oder 1 Jahr mit voller Lebensqualität gewonnen werden? Sind Behinderung und Lebensdauer beliebig kombinierbar?

- Wie werden Lebensjahre in verschiedenen Lebensabschnitten bei der Ermittlung der QALYs gewichtet? Ist es für den Wert eines QALYs gleichgültig, ob eine Therapie für ein Neugeborenes oder einen Rentner angewandt wird? Gibt es eine faire Anzahl an Lebensjahren, die jeder erreichen sollte (fair-innings-Ansatz)?

- Wie wichtig sind der Gesellschaft Unterschiede beispielsweise zwischen verschiedenen Formen der Behinderung? Sind Funktionseinschränkungen wichtiger als psychische Belastungen oder umgekehrt?

- Werden Krankheiten von Personengruppen mit Einkommen stärker gewichtet gegenüber Krankheiten von sozial schwachen Gruppen? Dies ist bei der Bemessung der indirekten Kosten (Arbeitsausfall) von Bedeutung. Oder sollen wohlhabende Bürger selbst vorsorgen, so dass schichtspezifische Prävalenzen in die Festlegung der Erstattung durch Sozialversicherungen einfließen sollen?

- Sollen Präferenzen der Krankheitsbekämpfung in der Bevölkerung mit einfließen in die Bewertung? Einige Krankheiten weisen sicher eine andere Wahrnehmung und Bewertung in der Bevölkerung auf als andere. Sollten Krankheiten stärker gewichtet werden, die besonders bedrohlich erscheinen?

- Wie soll zukünftiger Nutzen diskontiert werden? Vorgeschlagen werden Marktzinsen, also nach dem Investitionserfordernis der Behandlung, oder nach gefühlten subjektiven Abzinsungspräferenzen der Bevölkerung. Gibt es Schwellenwerte (beispielsweise für erreichte Lebensjahre), ab denen Diskontierungssätze zunehmen können?

- Wie wird Unsicherheit eingebaut? Sind Patienten risikoneutral bei der Entscheidung über Therapiealternativen?

Einige der Fragen erscheinen provokant in einem solidarischen Gesundheitssystem. Dennoch machen die Fragen deutlich, dass mit der Ermittlung von QALYs und deren Kosten auch Werturteile verbunden sind. Eine Festlegung für eine Gleichgewichtung von gewonnenen Lebensjahren unabhängig vom Alter der Patienten ist nun mal eine Entscheidung gegen eine unterschiedliche Gewichtung. Die zunächst für Allokationszwecke objektiv und wertneutral erscheinende Größe „Euro pro gewonnenem QALY“ ist es bei näherer Betrachtung nicht mehr.

Sculpher et al. (2005) und Kind (2005) gehen noch detaillierter auf

Literatur

Drummond MF, Barbieri M, Wong JB. Analytic choices in economic models of treatments for rheumatoid arthritis: What makes a difference? Med Decis Making. 2005;25(5):520-533.

Drummond MF, O'Brein B, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford; 1997.

Gyrd-Hansen D. Willingness to pay for a QALY: theoretical and methodological issues. Pharmacoeconomics. 2005;23(5):423-432.

IQWiG. Methodik für die Bewertung von Verhältnissen zwischen Nutzen und Kosten im System der deutschen gesetzlichen Krankenversicherung - Version 1.1. Fertigstellung am 09.10.2008. Köln. (http://www.iqwig.de/download/08-10-14_Methoden_Kosten-Nutzen-Bewertung_Version_1_1.pdf , Abfrage Nov 2008)

Kind P. Valuing health outcomes: ten questions for the insomniac health economist. In: Smith PC, Ginnelly L, Sculpher M (ed.). Health Policy and Economics. Open University Press: 2005: 42-63.

Roberts, Hsiao, Berman, Reich: Getting Health Care Reform Right. Oxford University Press: 2004.

Sculpher M, Claxton K, Akehurst R. It's just evaluation for decision making: recent developments in, and challenges for, cost-effectiveness research. In: Smith PC, Ginnelly L, Sculpher M (ed.). Health Policy and Economics. Open University Press: 2005: 8-41.

Steinbrook R. Saying No Isn't NICE — The Travails of Britain's National Institute for Health and Clinical Excellence. N Engl J Med 2008;359(19):1977-1981.

Young S, Godlee F. The BMJ Evidence Centre. BMJ 2008;337: a2438.

die technischen Details der Ermittlung von QALYs ein. Sie weisen korrekterweise darauf hin, dass verschiedene Instrumente der Nutzenmessung (insbesondere auch zur Ermittlung von QALYs) zu unterschiedlichen Ergebnissen und darauf basierenden Empfehlungen führen. Die Autoren sehen daher die Notwendigkeit einer Art Überführung, Kalibrierung oder auch Anpassung der Instrumente zur Messung der Lebensqualität untereinander. Inwieweit diese Verrechenbarkeit der Instrumente in regelmäßigen Zeitabständen selbst wieder neu kalibriert werden muss, bleibt dabei offen.

NICE hat daher Mindestanforderungen an die Ermittlung von QALYs gestellt. Demnach sind Index-Format, Kardinal-Skala von 0 bis 1 und die Ermittlung der Gewichte in relevanten Populationen vorgeschrieben. Hinzu kommt eine explizite Basierung auf Präferenzen, und zwar übergreifend über Krankheitsarten, also generisch. Um diese Vorgaben zu erfüllen, wird Standard Gamble (SG) meist als Goldstandard angesehen. Kind (2005) weist jedoch darauf hin, dass das Besondere an SG darin liegt, dass es keinen Standard der Durchführung gibt. Zudem können QALYs auch noch mit Time-Trade-Off Ansätzen (TTO) oder auch visuellen Analog-Skalen (VAS) berechnet werden. Sie können in enger Auslegung der Axiome von von-Neumann-Morgenstern jedoch nicht als utility bezeichnet werden. Die Literatur hierzu ist umfangreich und heterogen und trägt nicht dazu bei, die Vermittlung von CEAs in den politischen Prozess hinein zu vereinfachen.

Problematisch ist sicher, wenn zwei Verfahren, die das gleiche zu messen vorgeben, unterschiedliche Werte messen, etwa TTO und SG. In der Regel misst TTO geringere utility-Werte als SG. Liegen die Messwerte erst vor, kann aus ihnen nicht mehr erkannt werden, ob es sich um Nutzwerte im Sinne von von-Neumann-Morgenstern handelt. Nur das erklärt zugrunde liegende Messverfahren kann die Messwerte „veredeln". Kind (2005) fordert daher, dass ein einheitliches Verfahren ausgewählt wird, um QALY-Bewertungen vorzunehmen. Dazu gehört die Auswahl der Befragten, die Auswahl der Messung (etwa ein Verfahren des SG) und die Umrechnungsmodalitäten in Nutzwerte. Andere Verfahren sind dann nur noch gültig, wenn sie in diese Referenzwerte umgerechnet werden können. Bereits bei der Auswahl der Befragten räumt Kind allerdings ein, dass die Abbildung der „Normalbevölkerung" zur Ermittlung der Präferenzgewichte bereits Probleme bereiten kann (Gefängnisse, Streitkräfte etc.).

Wurde die Hürde der Messung innerhalb der CEA genommen, sind die Probleme noch nicht wesentlich geringer geworden. Die ermittelte ICER kann je nach eingeschlossenen Patienten, behandelndem Arzt, Wirkstoff, Region oder auch zufällig schwanken. Die ICER kann bei besonders betroffenen (selektierten) Patienten sehr günstig sein, dennoch wird eine Ausdehnung in die Fläche letztendlich unwirtschaftlich erscheinen.

Sculpher et al. (2005) fordert daher die Angabe einer Wahrscheinlichkeit, dass eine Intervention auch bei breiter Anwendung eine ICER erreicht, verbunden mit der Abschätzung, welche Ressourcen eine Regierung bei einer auf CEA basierenden Fehlentscheidung eventuell verloren geben muss. CEA eingesetzt als societal decision making, also rationale Verteilung eines vorgegeben Budget anhand von erreichbaren QALYs (aus exogen definierter gesellschaftlicher Sicht), halten die Autoren für zu einfach gedacht und gesellschaftlich nicht durchführbar.

Sie verweisen auch darauf, dass CEA immer nur eine Momentaufnahme liefern kann. Im Zeitablauf ändern sich die Erfassungs- und Auswertungsmethoden, der Zuschnitt der Perspektiven ändert sich (weshalb die Autoren grundsätzlich die gesellschaftliche Perspektive, und nicht etwa der Krankenkassen oder der betroffenen Patienten, vorschlagen), Technologie (=Therapien), Ressourcen und Information (=Evidenz) wandeln sich, und schließlich müssen die sich ändernden gesellschaftlichen Vorstellungen über Nutzen einbezogen werden. Ermöglicht eine Therapie beispielsweise erst das Erreichen eines gebärfähigen Alters, ist auch das Ereignis von Nachkommen und deren Nutzen (und wiederum deren Nachwuchses und Nutzen und so weiter) prinzipiell einzubeziehen, bis die Marginalisierung erreicht ist.

Sculpher et al. (2005) sind auch bei der Kostenerhebung überraschend kritisch. Eine Kostenerhebung auf Basis von Mittelwerten ist demnach nicht immer sinnvoll, da Kosten oft nicht normalverteilt sind, sondern schief. Es können extreme Ausreißer auftreten, prinzipiell ohne Limitierung nach oben. Eine Abschätzung des Einflusses der Therapie auf das Budget der Sozialversicherung (budget-impact Analyse) wird dadurch zumindest erschwert. Generell werden Kosten oftmals vereinfacht durch Erlöse abgeschätzt, was zu Verzerrungen führen kann. Produktivitätsverluste werden oftmals zwar erfasst (weil sie eher leicht zu erheben sind), haben jedoch höchst unterschiedliche Ausprägungen im internationalen Rahmen. Eine unkritische Übertragbarkeit der Ergebnisse aus dem Ausland verbietet sich daher.

Vor der politischen Umsetzung werden neue Therapien in CEAs zwar gegen bestehende Therapien (beziehungsweise sinnvolle Alternativmaßnahmen) verglichen und die marginalen Kostenzuwächse und Nutzengewinne gegenüber gestellt (ICER), jedoch kaum die Budgetobergrenzen thematisiert. Ein festes Gesamtbudget, welches jeder neuen Intervention die Bürde auferlegt, eine bestehende Intervention mit gleichem Budgetrahmen zu entfernen, wird nur selten angenommen. Vielmehr herrscht die Daumenregel vor, dass es ein dehnbareres Budget gibt, welches bis zu einem bestimmten Grenzwert (beispielsweise Euro/QALY) reichen kann. Diese Vorgehensweise vernachlässigt jedoch maßgeblich die Opportunitätskosten der neuen Therapie, denn andere (medizinische oder gesellschaftliche) Aktivitäten können nicht mehr vorgenommen werden.

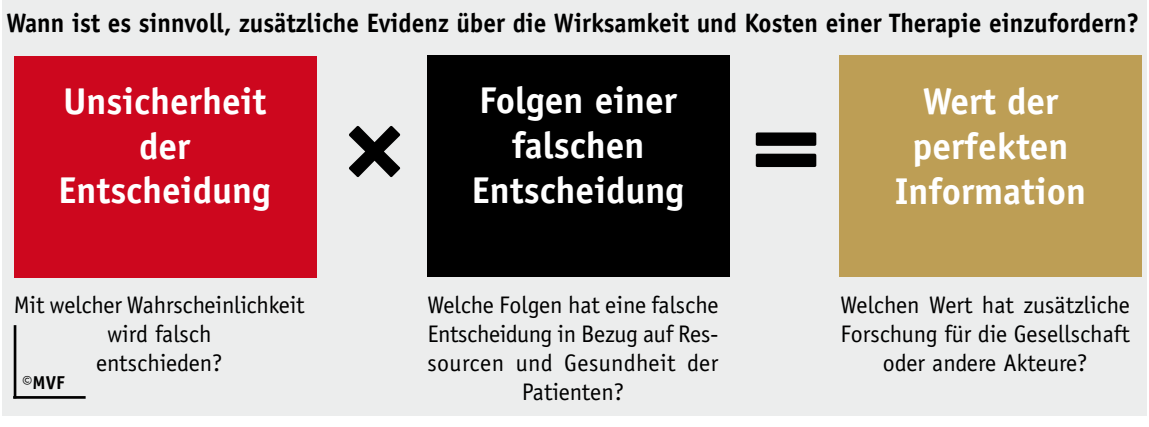


Abb. 2: Zusammenhang zwischen der Unsicherheit einer Entscheidung, den Folgen der falschen Entscheidungen und dem Wert der Information, die Entscheidung zu verbessern.

Um die „zu verdrängenden" Programme zu identifizieren, müssten auch für bestehende Programme deren marginale Kosten-Effektivität berechnet werden. In der Regel schwankt diese jedoch regional und zeitlich (wie auch für neue Programme) und ist nicht ansatzweise bekannt.

Die Kosten der Unsicherheit der Entscheidung, die von der Regierung auf der Basis von CEA getroffen werden, können interpretiert werden als der Wert perfekter Information (siehe Abb. 2). Neue Forschung (zur Verbesserung des Informationsstandes) sollte nicht teurer werden als dieser Wert. Daraus lässt sich beispielsweise auch eine Rangfolge für zukünftige Forschung ableiten. In einem weiteren Schritt können auf der Basis der „Kosten von Fehlentscheidungen" auch notwendige Studien entworfen werden in Bezug auf erforderlichen Stichprobenumfang, Zielgruppe, Follow-up Zeitrahmen und Wahl der Endpunkte.

Fazit

Sculpher et al. (2005) weisen auf den vergleichsweise geringen Einfluss hin, den CEA bei individuellen Entscheidungen auf Patientenebene haben, und auf den steigenden Einfluss auf Systemebene. Diese Aussage gilt insbesondere für England mit einer bereits 1999 installierten Institution wie NICE, jedoch auch für Deutschland. Eine Verankerung in der Wohlfahrtsökonomie wird kaum noch versucht. Die Rückbesinnung auf die Ursprünge der CEA als Entscheidungshilfe für ein konkretes Problem steht im Vordergrund.

Die Kritik an CEA darf nicht darüber hinweg täuschen, dass Entscheidungen im Gesundheitswesen nun mal getroffen werden müssen. Ein rein (subjektiv) utilitaristisch fundiertes Marktsystem wird in Deutschland, wie auch in den meisten anderen Industriestaaten, abgelehnt. Jedoch scheint sich die deutsche Diskussion um CEA derzeit in technischen Details zu erschöpfen. Das eigentliche Problem ist jedoch die bessere Einbindung der Studienergebnisse in Gesundheitspolitik und Versorgungspraxis (Young & Goodlee, 2008). Gute technische Qualität von CEA ist dabei nur ein Teilaspekt, womöglich nicht mal der ausschlaggebende. Seit die Ergebnisse von CEAs für die Pharmaindustrie millionenschwere Folgen haben, ist eine verblüffende Kreativität darum entbrannt, was denn eine gute technische Durchführung ist. Es wird spannend sein zu beobachten, ob die CEA dem Druck standhält, der auf ihr zunehmend lastet.

Aus meiner Sicht, und hier gehe ich über die gesichtete Literatur hinaus, können einige Tendenzen aufgezeigt werden, welchen Beitrag CEA zu leisten imstande ist. Erstens muss CEA für Transparenz sorgen und nicht für das Gegenteil. Dazu gehört der deutlichere Hinweis auf die Schwächen einer konkreten CEA. Alternativen der Modellierungen und deren Einfluss auf das Ergebnis sind zu dokumentieren. Gelangen CEA in den Ruf, dass mit ihnen prinzipiell jedes gewünschte Ergebnis gezeigt werden kann, wäre dies das Ende der Glaubwürdigkeit. Selbst Michael Drummond, der sicher unverdächtig ist, jedwede CEA zu verdammen, sieht eine erklärungsbedürftige Abweichung der Ergebnisse je nach Wahl des Modells und der Annahmen (Drummond et al., 2005). So ergaben seine Untersuchungen zum Einsatz von Infliximab bei der Behandlung von rheumatischer Arthritis je nach Modell Ergebnisse zwischen 20.000 und über 100.000 BPfund per QALY. Es ist unschwer vorauszusehen, welche Modellalternative ein Pharmahersteller präferieren würde. Es ist derzeit völlig unklar, wie solche Interessenkonflikte zwischen Hersteller, Gesellschaft und eventuell beauftragten Forschungseinrichtungen dokumentiert oder gar gelöst werden können. Die Bewährung von CEAs in der Praxis steht sicher noch bevor.

Zweitens sollten QALYs als Endpunkt nicht überbewertet werden. Es dürfte derzeit nahezu unmöglich sein, eine Studie ohne Messung der Lebensqualität aufzusetzen oder gar zu publizieren, sofern sich die Thematik auch nur im entferntesten dazu eignet. QALYs entstanden aus der onkologischen Versorgung, um zuvor formal kaum zu vereinbarende Ziele, Lebensqualität und Lebensdauer für Patienten anschaulich aufzubereiten. In der Onkologie sind signifikante Unterschiede die Regel, die verbleibende Lebensdauer oftmals relativ gering, und die Anwendung bleibt innerhalb eines medizinischen Kontextes und Patientenfällen. Die Ausweitung der QALY-Anwendung auf die prinzipielle Vergleichbarkeit aller Interventionen und Therapien kann hingegen kritisch gesehen werden. Die Ausdehnungen auf völlig unterschiedliche gesellschaftliche Kontexte, etwa Raucherprävention versus Schlaganfallversorgung, führt nicht dazu, die Diskussion tatsächlich zu vereinfachen.

Drittens lassen sich die Überlegungen der Kosten-Effektivität auch auf CEAs selbst anwenden. Ist absehbar, dass eine Studie keinerlei Relevanz haben wird, etwa weil das Studiendesign ungenügend ist oder das Gebiet bereits intensiv beforscht, kann auf eine Studie auch verzichtet werden. Die Konzentration der CEA auf die Bereiche, wo sie auch noch fühlbaren Erkenntnisgewinn bringen können, würde auch ihre Akzeptanz im politischen und gesellschaftlichen Umfeld stützen. <<

Abstract

In recent years, cost-effectiveness analyses (CEA) have increasingly become relevant in decision making in the health care sector. Yet, since legislation took up CEA as a tool for the evaluation of new therapies for reimbursement in the statutory health insurance (SHI), the strengths and liabilities of its methodology needs further investigation. In the UK and the U.S., simple decisions based on cost per QALY are politically and methodologically more and more perceived with skepticism. Moreover, the discussion around the priority of the subjective utility of the individual over the objective utility defined by a group of experts has not yet been resolved. Which of these two approaches to allocate resources in health care should be preferred is still an open question. Both approaches need further research, before they could give valuable support for decisions in health politics.

Keywords

cost-effectiveness analyses, welfare economics, health policy

Privatdozent Dr. rer. pol. Markus Lüngen

Kommissarischer Leiter des IGKE

Studium der Volkswirtschaft und Soziologie an der Universität zu Köln. Anschließend Arbeit als Referent für Krankenhausorganisation bei der Krankenhausgesellschaft Nordrhein-Westfalen. Seit 1999 Tätigkeit im Institut für Gesundheitsökonomie und Klinische Epidemiologie (IGKE) der Universität zu Köln und Habilitation im Fach Gesundheitsökonomie an der Wirtschafts- und Sozialwissenschaftlichen Fakultät. Seit 2005 Kommissarischer Leiter des IGKE und Vertreter der Professur Gesundheitsökonomie. Arbeitsschwerpunkte: Gesundheitspolitik, Finanzierungs- und Verteilungsfragen des Gesundheitswesens sowie Kosten-Nutzen-Analysen.
Kontakt: markus.luengen@uk-koeln.de

