

# Was genau ist eigentlich ein QALY?

Prof. Dr. med. Michael Schlander, M.B.A.



Quelle: privat

In vielen gesundheitsökonomischen Diskussionen taucht der Begriff des QALY auf. Manche Wissenschaftler halten das QALY für einen unverzichtbaren Standard, andere behaupten dagegen, die Benutzung von QALYs in ökonomischen Evaluationen sei entweder nicht in Übereinstimmung mit der ökonomischen Theorie oder auch, sie sei schlicht unethisch. Was also hat es mit dem ominösen Konstrukt auf sich?

» Schon seit 1968 arbeiten Gesundheitsökonomien mit QALYs. Seither wurde das Konzept zu einem der meistverbreiteten De-facto-Standards auf diesem Gebiet. Im Kern ist ein QALY, ausgeschrieben Quality-Adjusted-Life-Year, zunächst einmal nichts anderes als ein Maß für die Messung der gesundheitsbezogenen Ergebnisse („outcomes“) von medizinischen Maßnahmen. Ein wichtiges Ziel derer, die QALYs verwenden, ist die Vergleichbarkeit der Ergebnisse ganz unterschiedlicher Interventionen – etwa: Welchen gesundheitlichen Nutzen stiftet eine Appendektomie verglichen mit einer Chemotherapie bei einer malignen Erkrankung?

## ZWEI MESSUNGEN FÜR DEN ERFOLG

Eingreifende therapeutische Maßnahmen in der Onkologie zählen zu den beliebtesten Beispielen für die Begründung der Plausibilität des QALY-Konzepts: Die Messung einer Lebensverlängerung allein sagt nur die halbe Wahrheit über den Erfolg einer Therapie. Die zweite Hälfte liegt in der Lebensqualität der Patienten in dieser Zeitspanne. Im Fall einer aggressiven Chemotherapie kann diese erheblich eingeschränkt sein. Umgekehrt kann die Lebensqualität mit supportiven Maßnahmen (zum Beispiel der Gabe potenter Antiemetika) wiederum verbessert werden, ohne dass damit ein positiver Effekt auf die Überlebenszeit verbunden sein muss. Beide Effekte sind aber von Interesse bei der Bewertung medizinischer Interventionen. Das QALY kombiniert beide in einer einzigen Größe, indem die Lebenserwartung gewichtet wird mit einem Faktor  $q$ , der als Qualitätsindex auf

die Grenzwerte 1 für *vollständiges Wohlbefinden* und 0 für *Tod* normiert ist. Durch einfache Addition der mit dem jeweiligen Qualitätsmaß  $q$  bewerteten Zeiten lassen sich dann die gesundheitsbezogenen Effekte als QALYs bestimmen. In der Berechnung des gesundheitlichen Nutzens zählt daher ein in perfekter Gesundheit verbrachtes Lebensjahr vollständig, während sein rechnerischer Wert bei eingeschränkter Lebensqualität anhand des „Nutzwert“-Faktors  $q$ , der die gesundheitsbezogene Lebensqualität als Index abbildet, reduziert wird. Zum Beispiel beträgt der Faktor  $q$  Untersuchungen zufolge bei an Brustkrebs erkrankten Patientinnen beim ersten Rezidiv durchschnittlich 0,85, beim dritten dagegen nur noch 0,3. Nach einem schweren apoplektischen Insult mit erheblichen kognitiven und motorischen Defiziten wurde ein Wert von 0,03 beschrieben.

## „STANDARD GAMBLE“ MISST KARDINALE PRÄFERENZEN

Woher aber kommen diese Werte für  $q$ ? Immerhin stellen sie nicht weniger als einen Konversionsfaktor dar, der die proportionale Umrechnung von Lebensqualität in Lebenszeit und umgekehrt ermöglichen soll. Theoretisch müssen sie, damit die mathematischen Operationen der Addition von QALYs und des Vergleichs von QALY-Differenzen als Ergebnis unterschiedlicher Behandlungsstrategien erfüllt sind, mindestens auf dem Niveau einer Intervallskala (Ökonomen sagen dazu häufig „kardinal“) gemessen werden. Diesem Anspruch genügen prinzipiell die Methoden des Time-Trade-Off (TTO) und des Standard Gamble (SG). Das SG gilt als die klassische

Methode zur Messung kardinaler Präferenzen, da sie direkt auf der von den Spieltheoretikern John von Neumann und Oskar Morgenstern 1944 inaugurierten Erwartungsnutzentheorie gründet. Im Prinzip wird dabei jeder Gesundheitszustand, dessen Nutzwert („utility“)  $q$  bestimmt werden soll, mit der Teilnahme an einer „Lotterie“ verglichen, deren Eintrittswahrscheinlichkeiten für das Eintreten von zwei normierten Referenzzuständen (am einfachsten „Tod“ mit  $q = 0$  und „perfekte Lebensqualität“ mit  $q = 1$ ) bekannt sind. Es kann dann in diesem Beispiel die *Eintrittswahrscheinlichkeit*  $p$  für den sofortigen Tod ermittelt werden, bei welcher die Befragten „indifferent“ sind, das heißt, sich nicht mehr entscheiden können, ob sie den untersuchten Gesundheitszustand oder die Teilnahme an der Lotterie „präferieren“ würden. Daraus ergibt sich für den gesuchten Nutzwert  $q = (1-p)$ . Bei der TTO-Methode wird ebenfalls eine hypothetische Entscheidungssituation für die Ableitung des Wertes  $q$  benutzt (deshalb gelten sowohl SG als auch TTO als „choice-based methods“): Es wird im Prinzip so vorgegangen, dass diejenige Zeitspanne  $x$  an Lebensjahren in absoluter Gesundheit gesucht wird, bei welcher die Befragten wiederum indifferent werden gegenüber der *Lebenserwartung*  $t$  mit dem untersuchten Gesundheitszustand. Der zugehörige Nutzwert ergibt sich dann aus dem Quotienten beider Zeitspannen, also  $q = x/t$ .

## LEBENSQUALITÄT IN EINEM INDEX ABBILDEN

Da beide Verfahren methodisch aufwendig sind, wird in der Praxis gerne auf einfache-



Quelle: picture-alliance / obs

### Lebensqualität und Lebenserwartung – nur eine Multiplikationsaufgabe?

re Instrumente rekuriert, entweder auf „rating scales“ (Prinzip des „Fieberthermometers“) oder auf standardisierte Messinstrumente auf der Basis mehr oder weniger umfangreicher Fragenkataloge, welche gesundheitsbezogene Lebensqualität in einem Index abbilden. Dieser Index wird direkt – oder indirekt, nach einer normierten Transformation – überführt in Werte für  $q$ . Zu den gebräuchlichsten dieser Instrumente zählen der kanadische Health Utilities Index (HUI), die Quality-of-Well-Being-Skala (QWB), das Short Form 6D (SF-6D) und der europäische EuroQoL- oder EQ-5D-Index. Diese Instrumente unterscheiden sich untereinander erheblich bezüglich ihrer theoretischen Fundierung wie auch der berücksichtigten Dimensionen von Gesundheit. Es kann deshalb nicht ernsthaft überraschen, dass sie unter Umständen zu deutlich divergierenden Ergebnissen führen. Darüber hinaus gibt es systematische Unterschiede zwischen den Ansätzen SG, TTO und Rating Scales, welche sich ebenfalls (erklärbar!) in divergenten Werten für  $q$  niederschlagen. Eine verlässliche Umrechnung der auf unterschiedliche Weise gewonnenen Werte ist jedoch bisher nicht gelungen. Aus diesen Gründen ist es ratsam, sehr kleinen QALY-Differenzen (die sich in manchen Untersuchungen erst in der dritten Stelle hinter dem Komma zeigen) mit einer gesunden Skepsis zu begegnen. Dafür spricht auch die inzwischen vielfach belegte Insensitivität dieser Bewertungsmethode für kleine Differenzen. In solchen Fällen führen andere Evaluationsarten oft weiter.

### ADAPTIONSFÄHIGKEIT ALS NORMATIVES PROBLEM

Ein weiteres Problemfeld ergibt sich aus der Frage, aus welcher Gruppe die zu

Befragenden rekrutiert werden sollen. Gesunde bewerten hypothetische Gesundheitszustände anders (häufig schlechter) als betroffene Patienten. Hier zeigt sich die Fähigkeit des Menschen, sich an widrige Umstände anzupassen, mit großer Nachdrücklichkeit. Bei einmaliger Befragung (zumal bei Gesunden) besteht vor diesem Hintergrund generell das Risiko, eher die Schwellenangst vor einem ungünstigeren Gesundheitszustand zu messen als die damit tatsächlich verbundene Lebensqualität, die sich durch Adaptationsphänomene im Zeitverlauf deutlich verbessern kann. Es gibt deshalb zurzeit eine intensive akademische Diskussion

darüber, wie eine wirklich aussagekräftige, für daraus folgende Entscheidungen relevante Messung des Faktors  $q$  aussehen müsste.

### WEITERENTWICKLUNG ZWINGEND ERFORDERLICH

Feststehen dürfte zum gegenwärtigen Zeitpunkt jedenfalls, dass die derzeit üblichen Methoden aus den genannten und weiteren dem Modell inhärenten Gründen der Weiterentwicklung bedürfen. Bis dieses in einer überzeugenden Weise gelungen sein wird, stellt das QALY einen pragmatischen Ansatz dar, die gesundheitsbezogenen Effekte medizinischer Maßnahmen zu erfassen, dessen Eigenschaften und Limitationen zwischenzeitlich sehr gut verstanden werden. Es kann in ausgewählten Bereichen – nicht in allen (!), beispielsweise gibt es massive Probleme bei Anwendungen des Konzepts in der Pädiatrie – durchaus nützlich sein, solange man sich vor Überinterpretationen hütet und vor allem der (unter Gesundheitsökonomien vor allem in England verbreiteten) Versuchung widersteht, zugunsten des QALY auf die Heranziehung anderer, klinisch im Einzelfall aussagekräftigerer Endpunkte zu verzichten.

## Sommerakademie Gesundheitsökonomie (12.–16. Juni und 25.–29. September 2006)



Quelle: Gemeinde Kleinheubach

### Nähere Informationen:

Institut für Innovation & Evaluation im Gesundheitswesen (INNOVAL<sup>HC</sup>)

Frau Sigrid Steinebach

Telefon: 0 61 96 / 77 66-410

Telefax: 0 61 96 / 77 66-416

E-Mail: sigrid.steinebach@innoval-hc.de

Rathausplatz 12–14

65760 Eschborn

www.innoval-hc.com