

Zusammenfassende Gesundheitsmaße von Sterblichkeit und Krankheit: Der steinige Weg zwischen PYLL, YLD, DALY and HALE

G. Gmel^{1, 2, 3}
J. Rehm^{3, 4, 5}

Summary Measures of Public Health – the Rocky Road between PYLL, YLD, DALY and HALE

Zusammenfassung

Anliegen: Beschreibung von Gesamtmaßen der Gesundheit einer Population, wie verlorene Lebensjahre oder behinderungs-bereinigte Lebensjahre. **Methode:** Narrative Literaturdurchsicht. **Ergebnisse:** „Health-Expectancy“-Maße definieren die Lebens-erwartung einer Bevölkerung, „Health-Gap“-Maße die Differenz zwischen dem augenblicklichen Ist-Zustand und einem normati-ven Standard. Gesamtmaße der Gesundheit einer Bevölkerung schließen sowohl die Mortalität als auch die Morbidität einer Be-völkerung mit ein. **Schlussfolgerungen:** Trotz methodischer Schwächen informieren Gesamtmaße der Gesundheit politische Entscheidungsträger über eine verbesserte Allokation von Res-sourcen im Gesundheitswesen.

Schlüsselwörter

Gesundheitsmaße · PYLL · YLD · DALY · HALE

Abstract

Aims: To describe summary measures of public health such as life years lost or disability adjusted life years. **Methods:** Narra-tive literature review. **Results:** Health expectancy measures de-fine the life expectancy in a population, and health gap measures the difference between the actual situation and a normative standard. Summary measures of public health include both mor-tality and morbidity of a population. **Conclusions:** Despite some methodological problems summary measures of public health inform policy makers about a better allocation of resources in the public health sector.

Key words

Summary measure of public health · PYLL · YLD · DALY · HALE

Einleitung

Alkohol verursachte im Jahr 2000 4% der globalen Krankheits-belastung, in etwa genauso viel wie Tabak [1, 2]. Im gleichen Jahr wurden 3,2% aller Todesfälle durch Alkohol bedingt.

Wie kommt man zu solchen Aussagen? Dies ist das Thema dieser Veröffentlichung. Eine zentrale Rolle spielen dabei verschiedene Gesundheitsindikatoren wie verlorene Lebensjahre (YLL; years

life lost), verlorene Jahre durch Gesundheitseinschränkungen (YLD, years lost due to disability), qualitätsbereinigte Lebensjah-re (QUALY; quality adjusted life years) und behinderungsberei-nigte Lebensjahre (DALY, disability adjusted life years). Weiter-hin ist zum Verständnis dieser Indikatoren das Verständnis von technischen Fachtermini wie relatives Risiko (RR, relative risk), attributiver Anteil (AF, attributable fraction), kontrafaktisches Szenario (counterfactual scenario) notwendig. Hinter den meis-ten dieser Konzepte stehen komplexe mathematisch-epidemiolo-

Institutsangaben

¹ Swiss Institute for the Prevention of Alcohol and other Drug Problems (SIPA), Lausanne, Switzerland

² Alcohol Treatment Center, Lausanne University Hospital, Switzerland

³ Centre for Addiction und Mental Health, Toronto, Canada

⁴ Addiction Research Institute, Zurich, Switzerland

⁵ Public Health Sciences, University of Toronto, Toronto, Canada

Korrespondenzadresse

Gerhard Gmel · Swiss Institute for the Prevention of Alcohol and other Drug Problems · Avenue de Ruchonnet 14 · 1003 Lausanne, Schweiz · E-mail: ggmel@sfa-ispa.ch

Bibliografie

Suchttherapie 2006; 7: 143 – 153 © Georg Thieme Verlag KG Stuttgart · New York
DOI 10.1055/s-2006-927290
ISSN 1439-9903

logische Formeln, die an entsprechender Stelle nachgelesen werden können [3–6]. Die vorliegende Arbeit beschränkt sich darauf, die Konzepte intuitiv verständlich zu machen.

Ziel der vorliegenden Arbeit ist es also, zu beschreiben, welche Teilschritte notwendig sind, um so genannte „summary measures of public health“ (SMPH) zu berechnen, und welche Annahmen hinter der Berechnung der einzelnen Maße stehen. Als SMPH werden alle Maße bezeichnet, die das Ausmaß von Mortalität und Morbidität in einem Indikator zusammenfassen. Der Nutzen solcher Gesamtmaße liegt unter anderem in folgenden Aspekten begründet [7, 8]:

1. Vergleiche des Ausmaßes an Gesundheit und der Leistungsfähigkeit von Gesundheitssystemen zwischen verschiedenen Ländern.
2. Monitoring der Veränderungen von Gesundheitsaspekten innerhalb einer Gesellschaft.
3. Identifizierung und Quantifizierung von Gesundheitsunterschieden innerhalb eines Landes.
4. Einbeziehung nicht tödlicher Folgen bei der Evaluation von Gesundheitssystemen und der Gesundheit auf Bevölkerungsebene.
5. Bereitstellung von Informationen zur Festlegung von Prioritäten von Leistungen im Gesundheitssystem bzw. deren Planung.
6. Bereitstellung von Informationen, um Prioritäten in Forschung und Entwicklung festzulegen.
7. Analyse des Nutzens von Interventionen im Gesundheitsbereich, z. B. für Kosten-Nutzen-Analysen.

Wir basieren unsere Erläuterungen beispielhaft im Wesentlichen auf Alkohol- bzw. Tabakkonsum als Risikofaktoren. Vergleichende Aussagen wie „der Tabakgebrauch stellt den größten Risikofaktor für die Krankheitsbelastung insgesamt dar“ beruhen im Wesentlichen darauf, dass die Berechnungen getrennt für verschiedene Risikofaktoren durchgeführt und dann in Beziehung gesetzt werden. Im comparative risk assessment (CRA) im Rahmen der so genannten Global-Burden-of-Disease-Studie wurden 26 der bedeutsamsten Risikofaktoren untersucht, wie Untergewicht, Übergewicht, Cholesterin, Blutdruck, Früchte- und Gemüsekonsum, Alkohol-, Tabak- und illegaler Drogengebrauch, unsaubere Hygienebedingungen, Vitamin A-, Zink- und Eisenmangel oder körperliche Inaktivität [2]. Die generellen Berechnungen von „Summary Measures“ sind ähnlich für verschiedene Risikofaktoren und können in der Regel einfach von unseren Beispielen auf andere Risikofaktoren übertragen werden.

Wir beginnen unsere Erläuterungen mit relativ einfachen Gesundheitsindikatoren wie der Anzahl von Todesfällen und erweitern diese dann sukzessiv auf Gesamtmaße, die sowohl Indikatoren der Mortalität und der Morbidität beinhalten.

Ein erstes Beispiel für Gesundheitsindikatoren: Todesfälle

Der sicherlich einfachste Gesundheitsindikator beruht auf dem Gegenteil von Gesundheit, nämlich dem Tod. Man berechnet beispielsweise, wie viele Menschen in einem gewissen Zeitraum (z. B. in einem Jahr) infolge ihres Alkoholkonsums gestorben sind, und vergleicht diese Anzahl mit der Anzahl jener, die auf-

grund ihres Tabakkonsums oder ihres Übergewichts im gleichen Zeitraum gestorben sind. Da Menschen generell einmal sterben, bedeutet ein „alkoholbedingter“ Todesfall bei dieser Berechnung ein „vorzeitiger“ Todesfall, d. h. die jeweilige Person hätte ohne den „Risikofaktor“ länger gelebt. Das erste Problem an dieser scheinbar leichten Rechnung liegt in der Anwendung von Begriffen wie „infolge“ oder „aufgrund“. Gemeint ist damit „kausal“, also ursächlich, d. h. der Tod einer jeweiligen Person wäre nicht (zumindest nicht zu diesem frühen Zeitpunkt) eingetroffen, wenn die Person nicht Alkohol oder Tabak konsumiert hätte.

Wie jedoch kann man wissen, ob ein Todesfall „kausal durch Alkoholkonsum oder Rauchen bedingt“ war? Ist ein Raucher, der aufgrund eines Lungenkrebses im Alter von 60 Jahren stirbt, wegen des täglichen Konsums von einer Packung Zigaretten gestorben? Kennt nicht jeder in seiner Familie einen Rauchenden, der sogar noch mehr geraucht hat und im hohen Alter dann an einer Krankheit verstorben ist, die nach derzeitigem Wissensstand nicht mit Rauchen in Zusammenhang steht? Auch können Leute an Lungenkrebs erkranken, die nie in ihrem Leben geraucht haben. Ist der tödliche Autounfall mit 1,5 Promille im Blut immer kausal auf den Alkoholkonsum zurückzuführen? Ebenso wäre doch denkbar, dass die Verkehrssituation, bestimmte Straßenverhältnisse, technische Mängel am Fahrzeug oder unvorhergesehene Ereignisse auch dann zum tödlichen Unfall geführt haben könnten, wenn der Fahrer vorher keinen Alkohol getrunken hätte.

Viele Krankheiten, aber auch Unfälle haben nicht einen einzigen potenziell auslösenden kausalen Faktor, sondern mehrere. Beispielsweise werden mehr als 60 Krankheits- und Unfallkategorien [9, 10] durch Alkoholkonsum mitbedingt. Nur bei einigen Todesursachen, wie beispielsweise der alkoholischen Gastritis oder der alkoholischen Leberzirrhose, geht man per Definition davon aus, dass jeder Todesfall tatsächlich durch Alkohol verursacht worden ist. Wie sieht es bei den anderen Todesursachen aus? Wann zählt man denn beispielsweise den einen Krebstodesfall als einen „durch Alkohol“ verursachten und wann nicht?

Der entscheidende Punkt ist, dass man – anders als bei ausschließlich durch Alkohol verursachten Todesfällen (wie z. B. alkoholische Leberzirrhose, s. o.) nicht individuelle Todesfälle als durch Alkoholkonsum verursacht definieren kann, sondern nur entsprechende Anteile in der Bevölkerung, die statistisch bestimmt werden. Zentral dabei ist der Begriff der attributiven (zuschreibbaren) Anteile (AF, attributable fractions, auch etiologic fraction, population attributable fraction, preventable fraction genannt). Angenommen, der attributive Anteil von Alkohol für Brustkrebs bei Frauen in einer bestimmten Altersklasse in einem bestimmten Land beträgt 6%, so benötigt man keine individuelle Zuordnung mehr von Brustkrebstodesfällen zu den Kategorien „durch Alkoholkonsum verursacht“ oder „nicht verursacht“, sondern nur die Gesamtzahl an Todesfällen in diesem Land und der entsprechenden Altersgruppe.

Attributiver Anteil

Bei der Bestimmung attributiver Anteile unterscheidet man im Wesentlichen zwei Ansätze: einen direkten und einen indirekten. Die einfache Variante ist die direkte Zuschreibung. Beispiels-

weise stellen Polizei oder Feuerwehr fest, dass ein Brand durch die Zigarette eines im Bett eingeschlafenen Rauchers ausgelöst worden ist. Ihre Statistik weist beispielsweise aus, dass 20% (= attributiver Anteil) aller Brände durch Unachtsamkeit im Umgang mit Zigaretten oder Tabakwaren ausgelöst worden sind. Ähnliche Statistiken existieren auch für andere Todesursachen. Ein prominentes Beispiel sind durch Alkoholkonsum verursachte Verkehrstote. Solche direkten Zuordnungen sind nicht unproblematisch. Wer bestimmt beispielsweise, was ein durch Alkoholkonsum verursachter Verkehrsunfall ist? Was ist, wenn der Unfall auch passiert wäre, ohne dass der Fahrer Alkohol konsumiert hätte? Was ist mit nüchternen Personen, die durch einen alkoholisierten Fahrer tödlich verletzt worden sind? International hat sich durchgesetzt, erst bei einer Promillegrenze von 1,0 von Unfällen zu sprechen, die alkoholbedingt sind. Nach derzeitigen Erkenntnissen ist der Anteil der bei dieser Methode ausgeschlossenen Fälle größer als der Anteil der irrtümlich als alkoholbedingt klassifizierten.

Anders geht man bei der indirekten Methode zur Bestimmung attributiver Anteile vor. Zentral in diesem Zusammenhang sind Begriffe wie relatives Risiko, Prävalenz und Referenzgruppe. Der Begriff „Risiko“ ist in diesem Zusammenhang statistisch und nicht umgangssprachlich zu verstehen. Er bezeichnet eine Wahrscheinlichkeit. Anstelle zu fragen, wie wahrscheinlich es sei, eine 1 zu würfeln, könnte man auch fragen, wie hoch das Risiko sei, eine 1 zu würfeln. Formal kann die Wahrscheinlichkeit geschätzt werden durch die Häufigkeit des Eintretens eines Ereignisses geteilt durch die Summe von Ereignissen und Nichtereignissen. Findet man also bei 1000 Alkoholkonsumentinnen 110 Brustkrebserkrankungen (110 Ereignisse und 890 Nichtereignisse), so beträgt das Risiko 10% (= $110/(890 + 110) \times 100\%$). Beim relativen Risiko (RR) setzt man zwei Risiken in ein Verhältnis. Tritt in unserem Brustkrebsrisiko dieser nur bei 100 pro 1000 abstinent lebenden Frauen, so ist das relative Risiko $RR = 1,1$ (11%/10%). Man sagt auch, dass Alkoholkonsumentinnen ein 10% höheres Risiko im Vergleich zu abstinent lebenden Frauen haben. Wird das relative Risiko kleiner als 1, so spricht man von einem protektiven Effekt.

Bleiben wir beim obigen Beispiel von durch Alkoholkonsum bedingtem Brustkrebs. Wie viele Brustkrebserkrankungen wären in einer Bevölkerung dann rechnerisch durch Alkoholkonsum verursacht? Die Beantwortung dieser Frage hängt davon ab, wie groß der Anteil an Alkoholkonsumentinnen in dieser Bevölkerung ist. Nehmen wir vereinfachend im oben angegebenen Beispiel an, dass es genau 50% Alkoholkonsumentinnen gäbe, dann wären es bei 2000 Frauen 10 durch „Alkoholkonsum bedingte“ Fälle von Brustkrebs (110 versus 100) oder 4,76% (= 10 von 210) aller Brustkrebserkrankungen.

Epidemiologisch drückt sich dies in der folgenden Standardformel (z. B. [11, 12]) für Alkohol attributive Anteile (AAA) aus:

$$AAA = p_e \times (RR_e - 1) / [p_e \times (RR_e - 1) + 1],$$

wobei p_e die Prävalenz der „exponierten“ Gruppe (also im Beispiel der Alkohol konsumierenden Gruppe) ist und RR_e das relative Risiko dieser Gruppe im Vergleich zur nicht exponierten Gruppe. Im Rechenbeispiel ergäbe sich $0,5 \times 0,1 / (1 + 0,5 \times 0,1) = 0,05 / 1,05 = 0,0476$.

Das obige Prinzip lässt sich jetzt relativ einfach erweitern, wenn man nicht nur von einer einfachen Exposition (z. B. Alkoholkonsum versus Abstinenz) ausgeht, sondern von sich mit der Stärke der Exposition verändernden Risiken. Beispielsweise geht man davon aus, dass das Krebsrisiko mit der durchschnittlich konsumierten Menge ansteigt. In diesem Fall würde man verschiedene Gruppen definieren, z. B. Abstinente, Alkoholkonsum bis zu einem Glas täglich, bis zu 3 Gläsern täglich etc. Unter Verwendung der Prävalenz und des relativen Risikos (im Vergleich zur Referenzgruppe) für jedes dieser „Strata“ könnte man den attributiven Anteil (AA) gemäß folgender Formel berechnen:

$$AA = \sum p_e \times (RR_e - 1) / [\sum p_e \times (RR_e - 1) + 1].$$

Im Prinzip sind attributive Anteile also gewichtete Summen, wobei die Höhe des Risikos mit der Auftretenswahrscheinlichkeit der jeweiligen Exposition „gewichtet“ wird. Dies bedeutet, dass kleinere relative Risiken zu mehr attributiven Fällen führen können, wenn viele Personen einem solchen Risiko ausgesetzt sind, als höhere Risiken, die aber nur bei relativ wenigen Personen auftreten. Dies ist intuitiv einleuchtend. Auch wenn Alkoholkonsumierende von durchschnittlich 5 Gläsern am Tag oder mehr ein vielfach höheres Krebsrisiko aufweisen als durchschnittlich bis zu einem Glas täglich Konsumierende, so werden Letztere aufgrund ihrer höheren Anzahl in der Bevölkerung zu mehr alkoholbedingten Krebsfällen beitragen.

Wenngleich indirekt erhaltene attributive Anteile weniger von definitorischen Voraussetzungen abhängen wie die meisten direkt attributiven Anteile (z. B. alkoholbedingte Verkehrsunfälle als Definition einer Polizeistatistik), so hängt deren Validität von einer Vielzahl von Faktoren ab. Beispielsweise müssen relative Risiken, aber auch Prävalenzen unverzerrt geschätzt worden sein. In der Alkoholforschung weiß man zum Beispiel, dass methodische Faktoren, wie die Art des Messinstrumentes (Quantitäts-Frequenzmethode [QF] oder „graduierte Frequenzmethode“ [GF], z. B. [13–15]), ob Alkoholkonsum getränkespezifisch erfasst wird oder ob schriftlich oder telefonisch erhoben worden ist, zu erheblichen Unterschieden in der Schätzung der Prävalenz von Risikogruppen führen können. Darüber hinaus wird in etablierten Marktwirtschaften über Bevölkerungsbefragungen meist nur ein Teil (30–70%) der im jeweiligen Land verkauften Alkoholmenge abgedeckt, der Konsum also in der Regel drastisch unterschätzt. Auch wird die Exposition, also in unseren Beispielen der Konsum von Substanzen, in der Regel nur zu einem Messzeitpunkt erhoben und spiegelt so bestenfalls annähernd den Konsum über einen längeren Lebensabschnitt wider, da Individuen unterschiedlich starke Konsumphasen in verschiedenen Lebensabschnitten aufweisen. Bei vielen substanzbedingten Folgen ist aber die kumulierte Menge (z. B. die im Leben gerauchte Anzahl von Zigaretten, der Gesamtkonsum in Litern reinen Alkohols) ein stärkerer Prädiktor für Folgen als der Konsum zu einem bestimmten Zeitpunkt. Ebenso müssen bei der Schätzung relativer Risiken Probleme wie die Konfundierung durch andere Variablen beseitigt worden sein, um unverzerrte Schätzer zu erhalten.

Schließlich gibt es noch das Problem, dass relative Risiken von einer Population auf andere übertragen werden. Wenn beispielsweise für Deutschland der Anteil von alkoholbedingten Todesfällen bestimmt wird und man die entsprechenden Berechnungen

für die jeweiligen Krankheiten durchführt, greift man meist auf relative Risiken aus der internationalen Literatur zurück und setzt damit implizit voraus, dass diese über Länder hinweg stabil sind. Länderspezifische attributive Anteile entstehen so in der Regel nur durch länderspezifische Prävalenzen der Exposition nach Geschlechts- und Altersgruppen.

Doch selbst wenn sowohl Prävalenz als auch relative Risiken unverzerrt geschätzt worden sind, so ergeben sich Probleme der Interpretation. Eine Aussage, wie 6,98% aller Brustkrebs bei Frauen hätten verhindert werden können, wenn diese Frauen keinen Alkohol konsumiert hätten, setzt voraus, dass der Alkoholkonsum keinen Einfluss auf andere Krankheiten (competing risks) gehabt hätte. Dies ist aber unwahrscheinlich, da Alkoholkonsum eine Vielzahl von Gesundheitsfaktoren beeinflusst. Insgesamt lässt sich also sagen, dass für die indirekte Berechnung von attributiven Anteilen eine ganze Reihe von Annahmen getroffen wird, die sich auf die Ergebnisse und deren Interpretation auswirken können.

Kausalität

Attributive Anteile erlauben eine Abschätzung verschiedener Indikatoren der Krankheitsbelastung, sind also nicht nur auf Todesfälle anwendbar sondern ebenfalls auf Morbiditätsindikatoren wie Pflegetage in Krankenhäusern. Attributive Anteile stellen jedoch zunächst nur statistische Größen dar und sie lassen sich berechnen unabhängig davon, ob ein Risikofaktor kausal eine bestimmte Folge hervorruft oder nicht. Die Anwendung attributiver Anteile auf Todesursachen setzt jedoch voraus, dass diese durch den entsprechenden Risikofaktor mitverursacht werden. Ausreichende Evidenz für Kausalität schließt ein, dass ein Zusammenhang zwischen Exposition (z.B. Alkoholkonsum) und Folge (z.B. Krankheit oder Unfall) besteht, aber dass bei diesem Zusammenhang Zufallsergebnisse, Konfundierungen und andere plausible Alternativerklärungen ausgeschlossen werden können. Der rein statistische Zusammenhang zwischen Alkoholkonsum und Depression alleine besagt beispielsweise nicht, dass Alkoholkonsum vermehrt zu Depressionen führt. Die umgekehrte kausale Richtung „depressive Episoden führen zu verstärktem Alkoholkonsum“ oder Drittvariablen, die sowohl depressives Verhalten als auch Alkoholkonsum beeinflussen (z.B. genetische Veranlagung), könnten alternativ dem statistischen Zusammenhang zugrunde liegen. Die Bestimmung von Folgen, die kausal alkoholbedingt sind, erfordert häufig langjährige Forschung und die Abwägung verschiedener Kriterien (zum Beispiel der alkoholbedingten Depression, siehe z.B. [10]). Zurzeit werden über 60 Krankheits- und Unfallkategorien kausal als alkoholbedingt angesehen. Ob und welche Krankheitsbilder kausal dem Alkoholkonsum zugeschrieben werden, hat sich jedoch über die Zeit mit zunehmender wissenschaftlicher Erkenntnis verändert. So gab es beispielsweise 1995 noch keine ausreichende Evidenz [11], dass Brustkrebs bei Frauen kausal durch Alkoholkonsum mitbeeinflusst wird. Spätere Arbeiten machten diesen Zusammenhang jedoch plausibler (e.g. [16]; vgl. [17, 18]), sodass heutzutage davon ausgegangen wird, dass Alkoholkonsum kausal auf die Entstehung von Brustkrebs wirkt. Vergleiche von Krankheitsbelastungen verschiedener Risikofaktoren hängen also auch immer davon ab, welche Krankheiten in einer solchen vergleichenden Analyse miteingeschlossen worden sind.

Das kontrafaktische Szenario

Die Berechnung von alkohol- oder tabakbedingter Mortalität setzt immer ein alternatives, vergleichendes Szenario voraus (z. B. in der Wahl der Referenzgruppe bei der Bestimmung relativer Risiken). Anders formuliert stellt sich die Frage, welche Todesfälle unter welchen Umständen vermeidbar gewesen wären. Das gängige kontrafaktische Szenario in der Epidemiologie stellt die Frage [19]: Was wäre passiert, wenn es keine Exposition gegeben hätte? Es betrachtet also im Hinblick auf den Alkoholkonsum die totale Abstinenz einer Bevölkerung in der Vergangenheit als alternatives Szenario.

Dies bedeutet jedoch in der Regel, dass man mit einer rein hypothetischen und vermutlich nie realisierbaren Alternative rechnet [20]. Es ist zurzeit kaum denkbar, dass in Deutschland, Österreich oder der Schweiz kein Alkohol getrunken wird. Die Frage, wie viele Alkoholtote bei gesellschaftlicher Abstinenz also verhindert werden könnten, hat kaum einen Bezug zur Realität und informiert politische Entscheidungsträger nur unzureichend, insbesondere dann, wenn man aufgrund solcher Berechnungen die „einzusparenden Kosten“ als Anhaltspunkt für Investitionen betrachtet. Neuere Ansätze [21–23] postulieren deshalb, dass es sinnvoller wäre, statt attributiver Anteile vermeidbare (avoidable) Anteile zu schätzen bzw. den Effekt realisierbarer Reduktionen der Exposition zu schätzen [24]. Diese Ansätze werden zurzeit jedoch noch eher rudimentär umgesetzt, auch weil die Definition dessen, was denn „realistisch vermeidbar“ sei, recht komplex ist.

Zurück zum klassischen Ansatz. Dort geht es um die Frage der Quantifizierung der durch Exposition verursachten Mortalität und Krankheitsbelastung. Als Vergleichsgröße wird dann, wie oben ausgeführt, meist die Nullexposition angenommen, unter anderem, weil diese Exposition als mit der geringsten Krankheitsbelastung verbunden angesehen wird. Diese Annahme trifft aber nicht bei allen Expositionen zu. Sie mag vielleicht beim Rauchen gelten, da oft angenommen wird, dass jeglicher Tabakkonsum mit mehr negativen als positiven Konsequenzen assoziiert ist, obwohl es auch hier Fragezeichen gibt. Beim Alkoholkonsum liegen die Verhältnisse jedoch anders. Moderater Alkoholkonsum wirkt sich schützend auf eine Reihe von Krankheiten aus, prominent sind dabei ischämische Herzerkrankungen [10]. Studien zur sog. All-cause-Mortalität, also alle Todesursachen zusammengenommen, zeigen, dass moderater Alkoholkonsum insgesamt eher Todesfälle „verhindert“ [25], wobei „verhindern“ gleichzusetzen ist mit negativen attributiven Anteilen (prinzipiell können Todesfälle nicht verhindert, sondern der Eintritt eines Todesfalles nur herausgeschoben werden).

Für die Schweiz [26] wurde am Beispiel des Alkoholkonsums geschätzt, dass in einer Gesellschaft in der niemand Alkohol konsumiert (Abstinenz als kontrafaktisches Szenarium), im Vergleich zur augenblicklichen Konsumsituation etwa 730 mehr Todesfälle pro Jahr entstehen würden, Alkoholkonsum also 730 Todesfälle verhindert. Dagegen könnten in einer Gesellschaft, in der alle moderat konsumierten, im Vergleich zur augenblicklichen Situation 3460 Todesfälle „verhindert“ werden, es werden also in der augenblicklichen Situation 3460 Todesfälle verursacht. Die Wahl eines kontrafaktischen Szenarios beeinflusst somit entscheidend die Berechnung alkoholbedingter Todesfälle.

Zur Schweiz vergleichbare Ergebnisse wurden auch in anderen Ländern wie Kanada gefunden [27].

Von Todesfällen zu verlorenen Lebensjahren

Um einige wichtige Basiskonzepte zu erläutern, haben wir bisher auf Todesfälle fokussiert. Das reine Zählen von Todesfällen ist jedoch nur ein Aspekt der Quantifizierung von Mortalität und es weist einen entscheidenden Nachteil auf (vgl. [28]): Es berücksichtigt nicht den Zeitpunkt des Todes. Der Tod eines jungen Erwachsenen von 18 Jahren wird beim Auszählen von Todesfällen gleich behandelt wie der Tod eines 85-jährigen.

Ein Konzept, das dem unterschiedlichen Todesalter Rechnung trägt, ist dasjenige der „potenziell verlorenen Lebensjahre“ (PYLL; potential years of life lost). Bei der Beurteilung der Auswirkungen von Risikofaktoren ist wichtig zu wissen, dass PYLL und Anzahl von Toten unterschiedliche Indikatoren sind, die zu unterschiedlichen Aussagen kommen können. Im obigen Beispiel aus der Schweiz [26] wurde selbst bei einem kontrafaktischen Szenario der Abstinenz jährlich ein Verlust von mehr als 23 000 Lebensjahren geschätzt. Mit anderen Worten: Legt man den Indikator „Anzahl der Todesfälle“ zugrunde, ergeben sich mehr durch Alkoholkonsum verhinderte als durch diese Substanz verursachte Todesfälle, während man mit dem Indikator der potenziell verlorenen Lebensjahre zum Schluss kommt, dass Alkoholkonsum deutlich mehr verlorene Lebensjahre verursacht als verhindert. Dieser Unterschied ist, vereinfacht gesagt, dadurch bedingt, dass durch frühzeitige Mortalität, beispielsweise durch Verkehrsunfälle in jungen Jahren, mehr Lebensjahre durch Alkoholkonsum verloren gehen, als die lebensverlängernde Wirkung des Alkoholkonsum im höheren Alter bei ischämischen Herz-Kreislauf-Erkrankungen gewinnt.

Bei der Berechnung von PYLL versucht man zu quantifizieren, wie lange eine Person ohne die jeweilige Exposition gelebt hätte (deshalb „potenziell“ verlorene Lebensjahre, da man ja nicht wissen kann, wie lange jemand wirklich gelebt hätte). Die Berechnung von durch einen Risikofaktor (z. B. Alkoholkonsum) bedingten potenziell verlorenen Lebensjahren basiert dabei auf dem Zeitpunkt der ermittelten, durch Exposition bedingten Todesfälle (siehe oben). Dann werden von den jeweiligen Todeszeitpunkten die Jahre berechnet, welche die jeweilige Person jeweils noch durchschnittlich gelebt hätte, wenn sie nicht frühzeitig verstorben wäre (Details siehe unten).

Normativer Standard zur Berechnung verlorener Lebensjahre

Im Schweizer Beispiel wurde ein Alter von 74 Jahren als einfacher Grenzwert für die Berechnung verlorener Lebensjahre angenommen (der sog. normative Standard, [29]). Für die Berechnung verlorener Lebensjahre bedeutete dies, dass ein 18-jähriger Verkehrstoter 56 Jahre verliert, ein 70-jähriger jedoch nur 4 Jahre. Wenn PYLL auf einen solchen einfachen Standard basiert werden, bedeutet dies auch, dass Todesfälle nach 74 Jahren nichts mehr zur Berechnung der PYLL beigetragen haben, also rechnerisch ab 74 Jahren keine Lebensjahre mehr gewonnen oder verloren werden können. Prinzipiell kann man einen solchen Grundwert beliebig

wählen, und je nachdem, ob man den normativen Standard nach oben oder unten verschiebt, werden mehr oder weniger verlorene Lebensjahre geschätzt. In der Praxis hat sich dabei häufig durchgesetzt, dass die Lebenserwartung bei Geburt in einem Land als Grenzwert herangezogen wird.

Das oben ausgeführte Beispiel ist eine sehr einfache Art, PYLLs zu berechnen, die in der Literatur dafür kritisiert wurde, dass das Mortalitätsgeschehen im Alter nichts mehr zum Indikator beiträgt. Mit anderen Worten: In der zitierten schweizerischen Studie spielte es keine Rolle, ob jemand 75 Jahre oder 105 Jahre alt wurden. Im Allgemeinen sind solche Unterschiede aber sehr relevant, gerade im Vergleich von Ländern mit relativ hohen Lebenserwartungen (beispielsweise in Westeuropa, wo die durchschnittliche Lebenserwartung im Jahre 2005 fast 79 Jahre betrug, knapp 82 Jahre für Frauen und knapp 76 Jahre für Männer; siehe <http://www.cia.gov/cia/publications/factbook/geos/uk.html>).

Der klassische Ansatz, um auch das Mortalitätsgeschehen im Alter mit einzubeziehen, besteht deshalb darin, jeweils die bedingte Lebenserwartung zum Zeitpunkt des Todes in die Berechnung der PYLL einfließen zu lassen. Bedingte Lebenserwartung bedeutet hier die Lebenserwartung aller Personen des jeweiligen Alters, d. h. für diejenigen, die das entsprechende Alter bereits erreicht haben. Diese Lebenserwartung beträgt für die 74-jährigen in vielen Ländern durchaus mehr als 10 Jahre.

Dabei ist zu beachten, dass bei Vergleichen zwischen Ländern oder bei Schätzungen von Veränderungen innerhalb eines Landes über die Zeit der gleiche normative Standard verwendet werden muss, d. h. die bedingten Lebenserwartungen müssen gleich sein. Hier hat sich in der Praxis ein Standard herausgebildet, der von der hohen Lebenserwartung bei Geburt in Japan ausging und die bedingte Lebenserwartungen aufgrund von westlichen Gesellschaften nimmt (vgl. [30, 31]). In der Abb. 1 sind die entsprechenden bedingten Lebenserwartungen für diesen Standard nach Todesalter aufgeführt. Während bei Geburt für Männer eine Lebenserwartung von 80 Jahren angenommen wird und für Frauen von 82,5 Jahren und somit Totgeburten je nach Geschlecht entsprechend 80 bzw. 82,5 verlorene Lebensjahre zugewiesen bekommen, verringert sich die Anzahl verlorener Lebensjahre je nach Todesalter. Tritt der Tod nach 50 Jahren ein, ergeben sich noch 31 PYLL für Männer und 34 PYLL für Frauen, bei einem Todesalter von 70 13,6 PYLL für Männer und 16,2

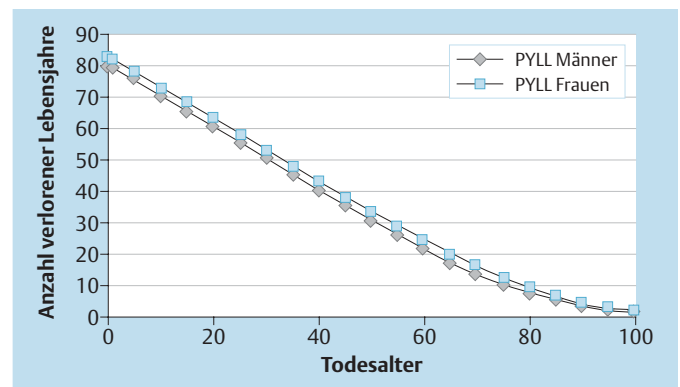


Abb. 1 Anzahl verlorener Lebensjahre in Abhängigkeit des Todesalters (Quelle: [7]).

PYLL für Frauen, bei einem Todesalter von 90 schließlich 3,5 PYLL für Männer und 4,3 PYLL für Frauen.

Dieser Standard scheint geeignet, PYLL in unterschiedlichen Ländern epidemiologisch zu vergleichen. Dabei gilt es noch zwei Varianten zu berücksichtigen: Wenn nicht epidemiologische, sondern ökonomische Sachverhalte im Mittelpunkt stehen, beispielsweise wenn verschiedene Gesundheitssysteme zu einem bestimmten Jahr verglichen werden sollen, so werden die verlorenen Lebensjahre üblicherweise diskontiert. Das bedeutet, dass die verlorenen Lebensjahre umso weniger gewichtig in die Vergleiche eingehen, je weiter sie vom Vergleichsjahr entfernt sind. Diese ist übliche ökonomische Praxis, nicht nur im Gesundheitsbereich, und solch diskontierten verlorenen Lebensjahre waren nicht nur die Basis für die GBD [6] und CRA [2], sondern auch für diverse Weltbankstudien (zuletzt [32]). Üblicherweise wird dabei ein Diskontsatz von 3% gewählt, d.h. mit jedem weiter entfernten Jahr wird ein Lebensjahr um 3% weniger stark gewichtet.

Zusätzlich zu den diskontieren PYLL gibt es noch eine andere Variante von PYLL, die in der Praxis Bedeutung erlangt hat. Diese Variante gewichtet nach Alter, d.h. ein verlorenes Lebensjahr im frühen Erwachsenenalter wird stärker gewichtet als ein verlorenes Lebensjahr im Alter. Technisch spricht man vom so genannten „age weighting“. Dieser Ansatz ist vom Ethischen her umstritten, weil er verlorene Lebensjahre bei unterschiedlichen Lebensabschnitten unterschiedlich gewichtet [29]. Die Verteidiger einer solchen Gewichtung führen an, dass in den meisten heutigen Gesellschaften de facto eine unterschiedliche Gewichtung stattfindet und dass age weighting deshalb die gesellschaftlichen Präferenzen besser widerspiegelt als eine ungewichtete Analyse.

Wie sich die unterschiedlichen Varianten auf die PYLL auswirken, kann Tab. 1 entnommen werden.

Von Indikatoren der Mortalität zu umfassenden Gesundheitsindikatoren

Wie im vorangegangenen Abschnitt gezeigt, liefert der Schritt vom reinen Zählen von Todesfällen hin zu verlorenen Lebensjahren eine zusätzliche Qualität, da Letztere auch das Lebensalter beim Eintritt des Todes berücksichtigen.

Mortalität spiegelt jedoch nur einen Teilaspekt der Gesundheit in Bevölkerungen wider. Darüber hinaus sind jedoch auch andere gesundheitsrelevante Folgen von Risikofaktoren wie Behinderung oder Krankheit von Bedeutung. Man stelle sich als ein Beispiel Verkehrsunfälle in zwei Ländern mit vergleichbarer Alkoholexposition vor. Bessere Gesundheitsversorgung, schnellere Wege zu Krankenhäusern, bessere Sicherheitsbedingungen im Straßenverkehr oder in den Fahrzeugen mögen in einem der beiden Länder mehr Todesfälle verhindern, aber zu einer größeren Belastung durch körperliche Beeinträchtigungen führen. Vergleiche zwischen Ländern, aber auch innerhalb eines Landes über die Zeit, die nur auf Todesfällen basierten, würden beispielsweise solche sozioökonomischen Ungleichheiten oder technologische Entwicklungen in der Gesundheits- und Verkehrswesen nicht geeignet berücksichtigen. Prinzipiell ist es auch wenig einleuchtend, nur die Todesursachen durch übermäßigen Alkoholkonsum zu berücksichtigen, jedoch nicht die Einschränkung der Lebensqualität durch alkoholbedingte Krankheiten, die für die Betroffenen und ihre Familien oft fast genauso bedeutsam sind. Für Planungen im Gesundheitssektor sind krankheitsbedingte Belas-

Tab. 1 Verlorene Lebensjahre mit und ohne Diskontierung und Altersgewichtung und der Standard West Level 26 (internationale Vergleichsnorm)

Altersgruppen	Durchschnittsalter bei Tod	Standard West Level 26		Diskontierung von 3% Altersgewichtung YLL [0.03, 1]		Diskontierung von 3% keine Altersgewichtung YLL [0.03, 0]		keine Diskontierung Altersgewichtung YLL [0.1]	
		Männer	Frauen	Männer	Frauen	Männer	Frauen	Männer	Frauen
0	0,1	79,94	82,43	33,11	33,22	30,30	30,52	85,90	87,20
1–4	2,6	77,77	80,28	35,16	35,28	30,10	30,33	85,54	86,84
5–9	7,5	72,89	75,47	37,21	37,35	29,59	29,87	82,26	83,59
10–14	12,5	67,91	70,51	37,30	37,47	28,99	29,31	76,75	78,09
15–19	17,5	62,93	65,55	36,02	36,22	28,29	28,67	69,96	71,32
20–24	22,5	57,95	60,63	33,84	34,08	27,47	27,93	62,55	63,93
25–29	27,5	52,99	55,72	31,11	31,39	26,53	27,07	54,96	56,36
30–34	32,5	48,04	50,83	28,08	28,40	25,44	26,08	47,50	48,93
35–39	37,5	43,10	45,96	24,91	25,30	24,18	24,94	40,39	41,85
40–44	42,5	38,20	41,13	21,74	22,19	22,74	23,63	33,75	35,24
45–49	47,5	33,38	36,36	18,63	19,16	21,09	22,13	27,68	29,19
50–54	52,5	28,66	31,68	15,65	16,26	19,22	20,45	22,21	23,73
55–59	57,5	24,07	27,10	12,83	13,52	17,14	18,55	17,38	18,88
60–64	62,5	19,65	22,64	10,19	10,97	14,85	16,43	13,16	14,62
65–69	67,5	15,54	18,32	7,80	8,60	12,42	14,09	9,60	10,93
70–74	72,5	11,87	14,24	5,71	6,45	9,99	11,59	6,72	7,82
75–79	77,5	8,81	10,59	4,00	4,59	7,74	9,07	4,53	5,32
80–84	82,5	6,34	7,56	2,68	3,09	5,78	6,76	2,93	3,44
85+	90,0	3,54	4,25	1,29	1,52	3,37	4,00	1,35	1,61

tungen deshalb von mindestens ebenso starker Bedeutung wie Todesfälle. Auch würde bei der ausschließlichen Betrachtung der Mortalität ein Ungleichgewicht in der relativen Bedeutung von Risikofaktoren entstehen. Risikofaktoren mit eher tödlichem Ausgang wie beim Tabakkonsum würden dann, wie es sich auch in der Vergangenheit der letzten Jahrzehnte gezeigt hat, im Vergleich zu Risikofaktoren wie Alkohol mit hoher Belastung durch Behinderung übergewichtet werden, auch wenn die meisten der tabakbedingten Todesfälle erst im hohem Alter entstehen. Ähnliches gilt für Krankheiten, die eher tödlich verlaufen, wie beispielsweise bestimmte Krebsarten, im Vergleich zu beispielsweise psychischen Erkrankungen wie Depression, die nicht tödlich verlaufen, aber über lange Zeit immer wieder zu deutlichen Einschränkungen der Aktivität und Lebensqualität führen. Historisch kann gezeigt werden, dass erst durch die Einführung von „summary measures of public health“ (SMPH) das Ausmaß der Belastung durch solche Krankheiten erkannt und bei der Gesundheitsplanung berücksichtigt worden ist.

SMPH kombinieren Informationen über Mortalität und nicht tödlichen Folgen für die Gesundheit (Morbidität) mit dem Zweck, die Gesundheit einer Population (z. B. eines Landes, Bundeslandes, Kantons etc.) in einem einzigen numerischen Wert zu repräsentieren [33].

Eine nahezu unüberschaubare Anzahl an solchen Gesamtmaßen ist entwickelt worden (für Überblicke siehe z.B. Murray et al. [29] oder Deebie [34]). Zurzeit haben sich im Wesentlichen die „disability adjusted life years“ (DALY) und die „healthy life expectancy“ (HALE) als Maße durchgesetzt, beispielsweise in der Global-Burden-of-Disease-Studie. Auf diese Maße wird im Folgenden fokussiert.

Generell wird unterschieden zwischen Maßen der „health expectancy“ (wie HALE) und Maßen des „Health Gap“ (z. B. DALY). Abb. 2 verdeutlicht die Unterschiede. Die Fläche A + B repräsentiert die Lebenserwartung zum Zeitpunkt der Geburt, wobei A die Zeit in perfekter Gesundheit (full health) repräsentiert und B die Zeit in weniger als perfekter Gesundheit. „Health-Expectancy“-Maße werden formal ausgedrückt wie folgt berechnet:

$$\text{health expectancy} = A + f(B),$$

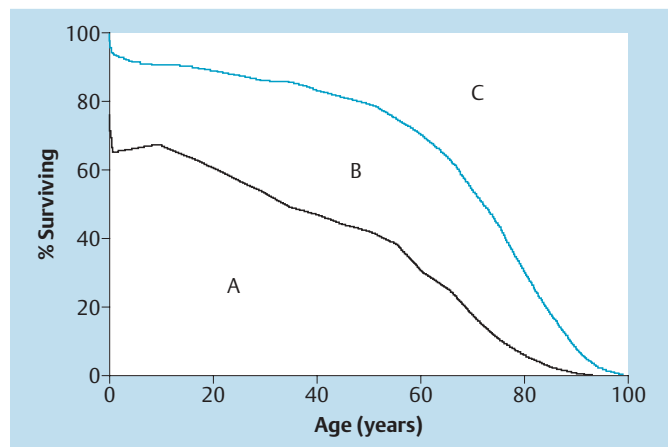


Abb. 2 Die Überlebenskurve (survivorship curve).

wobei $f(\cdot)$ eine Funktion ist, die verschiedene Gesundheitszustände unterschiedlich gewichtet. Dabei hat perfekte Gesundheit den Wert 1 und die Gewichte nehmen mit schlechterem Gesundheitszustand ab. Entscheidend ist dabei, wie diese Gewichte entwickelt werden. Darauf wird später eingegangen.

Im Gegensatz zu „Health Expectancy“ quantifizieren „Health-Gap“-Maße die Differenz zu einem normativen Standard (z. B. 100 Jahre als einfacher Standard in der Abb. 2), dargestellt durch die rechte, vertikale Begrenzung der Abb. 2. Potenziell verlorene Lebensjahre als Maße des „Mortality Gap“ würden dabei durch die Fläche C repräsentiert. Maße des „Health Gap“ gehen wiederum darüber hinaus, in dem sie auch die Zeit berücksichtigen, die nicht in perfekter Gesundheit verbracht wird. Formal ließe sich dies wie folgt ausdrücken:

$$\text{Health Gap} = C + g(B),$$

wobei $g(\cdot)$ wiederum eine Funktion ist, die Gesundheitszuständen schlechter als perfekte Gesundheit bestimmte Gewichte zuweist, wobei ein Gewicht von 1 bedeutet, dass die Zeit, die in einem solchen Gesundheitszustand verbracht worden ist, gleichbedeutend mit durch Mortalität verlorenen Lebensjahren ist.

Einfach ausgedrückt ist das Ziel von SMPH, Mortalität und Morbidität aufzusummieren und in einem einzigen Indikator auszudrücken, wobei unterschiedliche Schweregrade von Krankheit und Behinderung gewichtet werden. Ein verlorenes Lebensjahr durch vorzeitige Mortalität (Gewicht = 1) entspricht also beispielsweise 10 Jahren, die in Behinderung oder Krankheit gelebt worden sind, wenn der Schweregrad dieser Krankheit mit einem Gewicht von 0.1 versehen worden ist.

Health-Expectancy-Maße wie HALE messen dabei den augenblicklichen Ist-Zustand der durchschnittlichen Länge „gesunden Lebens“, also die Lebensdauer in perfekter Gesundheit. Health-Gap-Maße dagegen quantifizieren die Differenz zwischen dem augenblicklichen Ist-Zustand und dem, was sein könnte oder sollte.

Healthy life expectancy (HALE) und disability adjusted life years (DALY): Schätzung der in Behinderung und Krankheit verbrachten Lebensjahre (years lived with disability, YLD)

Das Kernstück zur Bestimmung von HALE oder DALY sind die sog. years lived with a disability (YLD). Disability meint dabei nicht etwa nur „Behinderung“, sondern alle Beeinträchtigungen (Krankheit und Behinderung), die nicht vollständiger Gesundheit entsprechen. Es ist vermutlich das schwierigste aller Teilmaße zur Bestimmung von Gesamtmaßen der Gesundheit einer Population und benötigt die Kenntnis verschiedener Bestimmungsstücke.

Die Grundformel zur Berechnung von YLD ist verhältnismäßig einfach:

$$YLD = I \times DW \times L,$$

wobei I die Anzahl der neuauftretenden Fälle ist, L die durchschnittliche Dauer der Krankheit/Behinderung und DW das sog.

„disability weight“, also der Schweregrad der jeweiligen Krankheit/Behinderung.

Für jede Krankheit und Behinderung werden YLD getrennt für unterschiedliche Altersgruppen und beide Geschlechter berechnet. Theoretisch würden also neben dem disability weight (weiter Erläuterungen zu disability weights siehe unten) für jede Krankheit nach Alter und Geschlecht in einem bestimmten Zeitabschnitt die Anzahl der Personen in dieser Altersgruppe, die Inzidenzrate, Remissionsrate (also die Rate der Genesung von einer Krankheit) sowie die Rate tödlicher Krankheits-/Behinderungsverläufe benötigt. Praktisch sind die unterschiedlichen Bestimmungsstücke für verschiedene Krankheiten/Behinderungen häufig nicht erhältlich, sodass man auf Prävalenzen, also in einem bestimmten Zeitabschnitt vorkommende Krankheits-/Behinderungskategorien zurückgreift, um dann mittels statistischer Formel die Inzidenz über einen bestimmten Zeitraum, meist ein Jahr, zu schätzen. Dazu ist dann minimal der Bevölkerungsaufbau der zu untersuchenden Regionen nach Geschlecht notwendig, die Prävalenz der zu untersuchenden Krankheit, ein Indikator für die bedingte Wahrscheinlichkeit, mit der entsprechenden Krankheit im untersuchten Zeitraum zu sterben, und die Remissionsrate, jeweils spezifisch nach Alter und Geschlecht.

Man kann YLD auch direkt aus der Prävalenz unter Verwendung von Daumenregeln der Beziehung zwischen Inzidenz und Prävalenz schätzen, um dabei folgende Formel verwenden (z. B. [35]):

$$YLD = \text{prev} \times DW$$

Dabei produzieren auf Inzidenz basierte und auf Prävalenz basierte Berechnungen ähnliche Ergebnisse, wenn sich der Gesundheitszustand in der Bevölkerung in der Vergangenheit bis zur Gegenwart nur wenig verändert hat (vgl. [36]). Prävalenzbasierte Ansätze sind dann zu bevorzugen, wenn lediglich die augenblickliche Krankheitsbelastung gemessen werden soll, inzidenzbasierte Ansätze jedoch, wenn man messen möchte, in welche Richtung sich eine Gesellschaft entwickelt, d. h. also bei der Abschätzung, wie der zukünftige Gesundheitszustand aussehen könnte [36]. Angesichts der häufig fehlenden Daten für Inzidenzen beruhen die meisten epidemiologischen Berechnungen zu YLD auf statistisch geschätzter Inzidenz.

Im Wesentlichen gibt die HALE die allgemeine Lebenserwartung abzüglich der YLD an. Hierbei ist jedoch zu berücksichtigen, dass YLD nicht über verschiedene Krankheiten/Behinderungen aufsummiert werden dürfen, da sich bestimmte Krankheiten wechselseitig bedingen und beeinflussen (Komorbidität). Eine Aufsummierung würde also die in nicht perfekter Gesundheit verbrachten Lebensjahre überschätzen. Entsprechende Methoden zur Korrektur sind entwickelt worden. Zur Bestimmung der HALE wird in der Regel auf die sog. Sullivan Methode [37] zurückgegriffen. Die WHO hat beispielweise HALE von 2000 für 191 Länder veröffentlicht [38].

DALY dagegen ist ein „Health-Gap“-Maß. Ein DALY kann verstanden werden als ein verlorenes gesundes Lebensjahr, wobei DALY als Gesamtmaß sowohl durch vorzeitigen Tod wie auch durch Krankheit verlorene Lebensjahre, gewichtet nach dem Schweregrad der Krankheit, einschließen. Die Krankheitsbelastung

(disease burden) ist die Differenz zwischen dem augenblicklichen Gesundheitszustand einer Population und einer globalen Referenzpopulation mit einer hohen HALE [39]. Einfach gesagt, sind DALY die Summe aus YLD und PYLL.

Disability weights

Um verlorene Lebensjahre durch Tod mit in Krankheit verbrachten Lebensjahren aufsummieren zu können, braucht es – wie bereits erwähnt – eine Gewichtung von Schweregraden verschiedener Krankheiten/Behinderungen. Vier Wochen mit einer Gipsschiene infolge eines Armbruches zu verbringen, hat eine unterschiedliche Bedeutung für das Individuum und die Gesellschaft als vier Wochen mit einem amputierten Arm zu leben oder – im Extremfall – 4 Wochen früher zu sterben. Disability weights zielen auf die Bestimmung solcher Gewichte ab. Sie liegen zwischen 0 (gesund) und 1 (tot).

Prinzipiell steht hinter einem solchen Konzept, dass zur Konstruktion eines SMPH Gesundheitszustände bewertet werden müssen bzw. der Gewinn (gain) beim Wechsel von einem schlechteren zu einem besseren Gesundheitszustand oder umgekehrt quantifiziert werden muss. Der Nutzen eines verbesserten Gesundheitszustand mag jedoch von einem betroffenen Individuum (individual utility) unterschiedlich bewertet werden als von der Gesellschaft (societal value). Auf die entsprechenden Unterschiede bei individueller und gesellschaftlicher Betrachtungsweise kann hier nicht eingegangen werden (vgl. dazu beispielsweise [40]). Wichtig ist jedoch, dass zur Bestimmung von disability weights immer auf Beurteilungen und Einschätzungen zurückgegriffen werden, in der Vergangenheit in der Regel auf die Einschätzungen von Expertenpanels wie in der Global-Burden-of-Disease-Studie [41]. Disability weights sind also keine „gottgegebenen“ Werte, sondern subjektive Einschätzungen, und wie Üstün, Rehm und Chatterji [42] zeigen konnten, variieren diese Einschätzungen von Schweregraden verschiedener Krankheiten trotz verhältnismäßig hoher allgemeiner Übereinstimmung in verschiedenen Ländern teilweise beträchtlich.

Es gibt verschiedene Bewertungsmethoden, anhand derer disability weights bestimmt werden, einschließlich der Verwendung von Ratingskalen (die bekannteste ist vermutlich die visual analogue scale, VAS, vgl. [43]) und sog. Trade-off- (Tauschgeschäft, Gegenleistung) Verfahren. Zu Letzteren gehören Methoden wie standard gamble (SG), time trade-off (TTO), willingness-to-pay (WTP) und person trade-off (PTO) (für einen Überblick vgl. [44]).

Bei der Verwendung der VAS bekommen die Urteiler z. B. eine verankerte Skala von 0 (gesund) bis 100 (tot) vorgelegt und müssen verschiedene Krankheitsbilder in dieser Skala einordnen. Vorteil dieses Skalierungsverfahrens ist die leichte Anwendbarkeit, ein Nachteil ist, dass insgesamt erfahrungsgemäss nur ordinale Urteile (A ist besser als B) erreicht werden, für die Bestimmung von disability weights aber Gewichte auf einer Verhältnisskala notwendig sind (d. h. Aussagen wie „die Krankheit A führt zu doppelt so starker Behinderung wie B“). VAS schließt auch keinen trade-off mit ein und führt so häufig zur Überschätzung des Schweregrades der Krankheit.

Bei Trade-off-Verfahren muss der Urteiler hypothetisch etwas Wertvolles (valuable) im Austausch dafür hergeben, dass eine Verbesserung des zu beurteilenden Gesundheitszustandes im Vergleich zu perfekter Gesundheit erreicht wird. Diese Kosten werden als Eichung (calibration) bezeichnet.

- Im SG muss der Urteiler angeben, welches Risiko eines potenziell möglichen sofortigen Todes er eingehen würde (z. B. das Risiko, bei einer Augenoperation zu sterben), um von dem zu bewertenden Krankheitszustand (hier als Beispiel Erblindung) gesunden zu können. Die Annahme ist, dass je stärker die Krankheit/Behinderung eingeschätzt wird, ein umso höheres hypothetisches Risiko eingegangen wird.
- Im TTO besteht die Eichung aus Lebensjahren, die ein Urteiler opfern würde, um von der Krankheit gesundet zu leben. Beispielsweise könnte ein Urteiler angeben, er würde in Kauf nehmen, 5 Jahre früher zu sterben, wenn er von einer Erblindung für den Rest des dann noch verbleibenden Lebens geheilt werden könnte. Annahme ist, dass je schwerer die Krankheit beurteilt wird, desto mehr zu opfernden Lebensjahre angegeben werden.
- Im WTP wird der Urteiler gefragt, wie viel er bereit wäre, für seine Gesundheit zu bezahlen; beispielsweise wie viel er für ein potenzielles Mittel bezahlen würde, wenn ihn dieses für ein Jahr von der Krankheit heilen könnte.
- Im PTO müssen Urteiler angeben, wie viele Personen mit der zu beurteilenden Krankheit sie als äquivalent für eine festgelegte Anzahl an Personen in perfekter Gesundheit ansehen.

Allen Trade-off-Ansätzen ist gemein, dass das Leben in Krankheit oder Behinderung per Definition als weniger wertvoll angesehen wird als jenes in perfekter Gesundheit, was von verschiedenen Forschern als unethisch betrachtet wird (z. B. [45]). In der Global-Burden-of-Disease-Studie wurden im Wesentlichen PTO-Methoden eingesetzt, weil für die Bestimmung der DALY die gesellschaftliche Perspektive relevant war.

Das folgende Beispiel soll einen PTO-Ansatz verdeutlichen (vgl. [43]). Man bittet den Urteiler, sich Folgendes zu vergegenwärtigen. Eine Intervention A erlaubt es, das Leben von 1000 vollständig gesunden Menschen um ein Jahr zu verlängern. Ohne Intervention A würden sie sofort sterben. Es gibt aber eine zweite mögliche Intervention B, die es Blinden (hier als Beispiel des zu evaluierenden Gesundheits-/Behinderungszustandes) erlauben würde, ein Jahr länger leben zu können (allerdings weiterhin blind). Wiederum würden sie ohne Intervention sofort sterben. Leider ist es wegen begrenzter finanzieller Mittel nur möglich, alternativ eine der Interventionen durchzuführen. Die Aufgabe des Urteilers ist es nun, die Anzahl an Blinden zu bestimmen, bei der er Intervention B anstelle von A wählen würde. Die gewählte Anzahl Blinder muss dabei größer als 1000 sein. Angenommen, ein Urteiler wählte ein trade-off von 1250 Personen, so wäre die Bewertung von Blindheit $1000/1250 = 0,8$ und das disability weight wäre 0,2. Neben ethischer Bedenken hat sich die Operationalisierung von Trade-off-Methoden also äußerst schwierig und schwer verständlich erwiesen, sodass die Aufgabenstellungen eher Intelligenztests ähneln [46] und Schätzungen vermutlich in Richtung auf jene besser Gebildeter verzerrt sind. Ähnliche Probleme zeigen sich auch bei anderen Trade-off-Methoden. So sind beispielsweise reichere Personen im WTP-Ansatz geneigt, mehr zu bezahlen als ärmere.

Tab. 2 gibt einen Überblick zu disability weights von Kernindikatoren der Global-Burden-of-Disease-Studie 1990.

Tab. 2 Durchschnittliche Gewichte für Krankheitsschweregrade (disability weights) für 22 Kernindikatoren (Quelle Murray and Lopez, 1996, S. 39 [3])

Bedingung	durchschnittliches disability weight
Weißfleckenkrankheit (vitiligo) im Gesicht	0,020
Abweichung um mehr als zwei Standardabweichungen vom Durchschnittsgewicht/-größe	0,024
Durchfall (Diarrhö)	0,066
schwere Halsschmerzen	0,077
schwere Blutarmut (Anämie)	0,111
eingeschränkte Bewegungsfreiheit im Gipsverband	0,136
Unfruchtbarkeit	0,191
Erektionsstörungen	0,195
rheumatoide Arthritis	0,209
Angina	0,223
Amputationen unterhalb des Knies	0,281
Taubheit	0,333
milde Formen geistiger Entwicklungsverzögerungen	0,361
rektovaginale Fistel	0,373
Down-Syndrom ohne Herzfehler	0,407
unipolare Depression	0,619
Blindheit	0,624
Querschnittslähmung	0,671
aktive Psychose	0,722
schwere Migräne	0,738
Demenz	0,762
Quadriplegie	0,895

Abschließende Bemerkungen

Die vorliegende Arbeit hat versucht, die Hintergründe und Annahmen bei der Berechnung von summary measures of public health (SMPH) aufzuzeigen. Dabei wurde auch auf die Komplexität der Berechnung und die potenziellen Schwierigkeiten bei der Interpretation und Anwendung solcher Maße aufgezeigt. Das Vorhandensein von Schwierigkeiten bedeutet jedoch nicht, dass auf die Anwendung solcher Maße verzichtet werden sollte. Sicherlich sind auf der einen Seite zusammenfassende Maße aussagekräftiger als das reine Zählen von Todesfällen oder reine Mortalitätsmaße wie „potenziell verlorene Lebensjahre“ (PYLL; potential years of life lost), welche die Morbidität unberücksichtigt lassen. Auf der anderen Seite spiegeln Absolutaussagen wie „weltweit werden jährlich 58 323 000 behinderungsbereinigte Lebensjahre (DALY, disability adjusted life years) verloren“ angesichts der fehlenden Datengrundlage für eine Vielzahl der benötigten Bestimmungstücke (z. B. Krankheitsprävalenzen oder -inzidenzen), die Beliebigkeit der Verwendung eines normativen Standards, z. B. bei der Bestimmung der „healthy life expectancy“ (HALE), oder die Variabilität in der Bestimmung von disability weights in Abhängigkeit unterschiedlicher Methoden (verschiedene time-trade offs oder visual analogue scale [VAS]) vermutlich nur eine Pseudogenauigkeit vor. Absolutaussagen auf globaler Ebene sind aber auch nicht das primäre Ziel der Gesund-

heitsberichterstattung der WHO. Es geht vielmehr um vergleichende Aussagen zwischen Ländern und Gesundheitssystemen, und hierzu wurde in den letzten Jahren ein Instrumentarium geschaffen, das zum ersten Mal in der Geschichte der Epidemiologie Vergleiche zulässt, die über bloße Mortalitätsunterschiede hinausgehen.

Deshalb glauben wir, dass SMPH einen wichtigen Beitrag geleistet haben, um aussagekräftige Unterschiede zwischen Ländern, Veränderungen innerhalb von Ländern über die Zeit oder der relativen Relevanz verschiedener Risikofaktoren zu einem gegebenen Zeitpunkt (z. B. Alkohol versus illegale Drogen) zu erfassen, und somit zu einer evidenzbasierten Gestaltung der Gesundheitspolitik beitragen. Dessen ungeachtet können und sollten solche Maße methodologisch weiterentwickelt werden, um bestehende Schwächen zu überwinden.

Literatur

- 1 World Health Organization (WHO). The World Health Report 2002 – Reducing Risks, Promoting Healthy Life. Geneva: World Health Organization (WHO), 2002
- 2 Ezzati M, Lopez AD, Rodgers A et al. Comparative Quantification of Health Risks. Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. Geneva: World Health Organization (WHO), 2004
- 3 Murray CJL, Lopez AD. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020. Cambridge, MA: Harvard School of Public Health on behalf of the World Health Organization and the World Bank, 1996
- 4 Murray CJL, Salomon JA, Mathers CD et al. Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- 5 Mathers CD, Bernard C, Moesgaard Iburg K et al. Global Burden of Disease in 2002: Data Sources, Methods and Results. Global Programme on Evidence for Health Policy Discussion Paper No. 54 (revised February 2004). Geneva: World Health Organization (WHO), 2003
- 6 Lopez AD, Mathers CD, Ezzati M et al. Global Burden of Disease and Risk Factors New York, NY: Oxford University Press & The World Bank, 2006
- 7 Murray CJL, Salomon JA, Mathers CD. A critical examination of summary measures of population health. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- 8 Richardson J. Evaluation summary measures of population health. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- 9 Gutjahr E, Gmel G. Defining alcohol-related fatal medical conditions for social-cost studies in Western societies: an update of the epidemiological evidence. *J Subst Abuse* 2001; 13 (3): 239–264
- 10 Rehm J, Room R, Monteiro MG et al. Alcohol use. In: Ezzati M, Lopez AD, Rodgers A, Murray CJL (Hrsg). Comparative Quantification of Health Risks Global and Regional Burden of Disease Attributable to Selected Major Risk Factors (Vol. 1). Geneva: World Health Organization (WHO), 2004
- 11 English DR, Holman CDA, Milne E et al. The Quantification of Drug Caused Morbidity and Mortality in Australia, 1992. Canberra: Commonwealth Department of Human Services and Health, 1995
- 12 Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. *Am J Epidemiol* 1974; 99 (5): 325–332
- 13 Gmel G, Rehm J. Measuring alcohol consumption. *Contemp Drug Probl* 2004; 31 (Fall): 467–540
- 14 Feunekes GJ, van't Veer P, van Staveren WA et al. Alcohol intake assessment: the sober facts. *Am J Epidemiol* 1999; 150 (1): 105–112
- 15 Rehm J. Measuring quantity, frequency, and volume of drinking. *Alcohol Clin Exp Res* 1998; 22 (2 Suppl): 4S–14S
- 16 Smith-Warner SA, Spiegelman D, Yaun SS et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. *JAMA* 1998; 279 (7): 535–540
- 17 Gutjahr E, Gmel G. Die sozialen Kosten des Alkoholkonsums in der Schweiz: Epidemiologische Grundlagen 1995–1998 (Forschungsbericht 36). Lausanne: Schweizerische Fachstelle für Alkohol- und andere Drogenprobleme (SFA), 2001
- 18 Ellison RC, Zhang Y, McLennan CE et al. Exploring the relation of alcohol consumption to risk of breast cancer. *Am J Epidemiol* 2001; 154 (8): 740–747
- 19 Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol* 2002; 31 (2): 422–429
- 20 Single E, Collins DJ, Easton B et al. International Guidelines for Estimating the Costs of Substance Abuse. Geneva: World Health Organization (WHO), 2003
- 21 Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fractions. *Am J Epidemiol* 1988; 128 (6): 1185–1197
- 22 Murray CJL, Ezzati M, Lopez AD et al. Comparative quantification of health risks conceptual framework and methodological issues. *Popul Health Metrics* 2003; 1 (1): 1–20
- 23 Rehm J, Taylor B, Patra J et al. Avoidable burden of disease: conceptual and methodological issues in substance abuse epidemiology. *Int J Meth Psych Res* in press
- 24 Collins DJ, Lapsley HM, Brochu S et al. International Guidelines for the Estimation of the Avoidable Costs of Substance Abuse. Ottawa: Health Canada, 2006
- 25 Rehm J, Gutjahr E, Gmel G. Alcohol and all-cause mortality: a pooled analysis. *Contemp Drug Probl* 2001; 28 (3): 337–361
- 26 Gutjahr E, Gmel G. Association of alcohol consumption to mortality and person-years of life lost in Switzerland – measuring the impact of some methodological options. *Eur J Epidemiol* 2005; 20: 37–47
- 27 Rehm J, Giesbrecht N, Patra J et al. Estimating chronic-disease deaths and hospitalizations due to alcohol use in Canada 2002 – Implications for policy and prevention strategies. *Prev Chronic Dis* in press
- 28 Wolfson MC. On the uses of summary measures of population health In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- 29 Murray CJL, Mathers CD, Salomon JA et al. Health gaps: an overview and critical appraisal. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- 30 Coale AJ, Demeny P. Regional Model Life Tables and Stable Population. New York, NY: Academic Press, Inc, 1983
- 31 Coale AJ, Guo G. Revised regional model life tables at very low levels of mortality. *Popul Index* 1989; 55 (4): 613–643
- 32 Jamison DT, Breman JG, Measham AR et al. Disease Control Priorities in Developing Countries (2nd ed) New York, NY: Oxford University Press & The World Bank, 2006
- 33 Field MJ, Gold MR. Summarizing Population Health: Directions for the Development and Application of Population Metrics. Washington, DC: National Academy Press, 1998
- 34 Deeble J. Resource Allocation in Public Health: An Economic Approach (A background discussion paper for the National Public Health Partnership) Melbourne: National Public Health Partnership, 2000
- 35 Mathers CD, Murray CJL, Salomon JA et al. Estimates of Healthy Life Expectancy for 191 Countries in the Year 2000: Methods and Results (GPE Discussion Paper No. 38). Geneva: World Health Organization (WHO), 2001
- 36 Barendregt JJ. Incidence- and prevalence-based SMPH: making the twain meet. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- 37 Sullivan DF. A single index of mortality and morbidity. *HSMHA Health Rep* 1971; 86 (4): 347–354
- 38 World Health Organization (WHO). The World Health Report 2001 – Mental Health: New Understanding, New Hope. Geneva: WHO, 2001
- 39 World Health Organization (WHO). The World Health Report 2003 – Shaping the Future. Geneva: World Health Organization (WHO), 2003
- 40 Nord E. My goodness – and yours: a history, and some possible futures, of DALY meanings and valuation procedures. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). Summary Measures of

- Population Health: Concepts, Ethics, Measurement and Applications. Geneva: World Health Organization (WHO), 2002
- ⁴¹ Murray CJL. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull World Health Organ* 1994; 72 (3): 429 – 445
- ⁴² Üstün B, Rehm J, Chatterji S. Are disability weights universal? Ranking of the disabling effects of different health conditions in 14 countries by different informants. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. Geneva: World Health Organization (WHO), 2002
- ⁴³ Stouthard MEA, Essink-Bot ML, Bonsel GJ et al. *Disability Weights for Diseases in the Netherlands*. Rotterdam: Department of Public Health, Erasmus University Rotterdam, 1997
- ⁴⁴ Essink-Bot ML, Bonsel GJ. How to derive disability weights. In: Murray CJL, Salomon JA, Mathers CD, Lopez AD (Hrsg). *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. Geneva: World Health Organization (WHO), 2002
- ⁴⁵ Arnesen T, Nord E. The value of DALY life: problems with ethics and validity of disability adjusted life years. *Br Med J* 1999; 319 (7222): 1423 – 1425
- ⁴⁶ Dolan P, Gudex C, Kind P et al. The time trade-off method: results from a general population study. *Health Econ* 1996; 5 (2): 141 – 154