

Wie kommuniziert man Risiken?

Gerd Gigerenzer

Am Anfang des 20. Jahrhunderts hat der Vater der modernen Science Fiction, Herbert George Wells, in seinen politischen Schriften Folgendes prophezeit: Wenn wir mündige Bürger in einer modernen technologischen Gesellschaft möchten, dann müssen wir ihnen drei Dinge beibringen: lesen, schreiben und – statistisches Denken. Zu Beginn des 21. Jahrhunderts hat in unserem Land praktisch jeder lesen und schreiben gelernt – nicht aber statistisches Denken. Mit statistischem Denken meine ich den rationalen Umgang mit Unsicherheiten und Risiken. Die Kunst, Risiken zu verstehen und klar zu kommunizieren, ist beispielsweise eine notwendige Voraussetzung um „shared decision making“ oder „informed consent“ in der Arzt-Patient-Interaktion überhaupt möglich zu machen.

Die mangelnde Fähigkeit zum statistischen Denken liegt nicht allein in der „Zahlenblindheit“ mancher Ärzte und Patienten, sondern in der *Form*, in der diese Information kommuniziert wird. Heute Abend spreche ich über drei Formen der Risikokommunikation, die regelmäßig Missverständnisse und Verwirrung erzeugen. Dann zeige ich jeweils alternative Methoden auf, welche Klarheit ermöglichen. Beginnen werde ich mit der Illusion von Gewissheit, einer mentalen Haltung, welche die Kommunikation von Risiken gar nicht notwendig erscheinen lässt, da sie stattdessen Sicherheiten vorspiegelt.

Die Illusion von Gewissheit

Von Benjamin Franklin stammt der Spruch: „In dieser Welt ist nichts gewiss, außer dem Tod und den Steuern.“ Diese Lektion haben wir in unserer Gesellschaft nicht immer gelernt. Nehmen wir einmal HIV-Tests als Beispiel, welche zu den zuverlässigsten Tests gehören, die wir haben. Diese erzeugen, wie alle Tests, einen gewissen Anteil von falsch-positiven und falsch-negativen Ergebnissen, der auf etwa 0,01 und 0,1 Prozent geschätzt wird (bezogen auf wiederholte ELISAs und einem Western-Blot-Test, eine Blutprobe). In anderen Worten, die Spezifität und die Sensitivität des Tests sind etwa 99,99 und 99,9 Prozent (Gigerenzer, Hoffrage & Ebert, 1998). Die Prävalenz des Virus bei deutschen Männern, welche kein Risikoverhalten praktizieren (z.B. nicht intravenös Drogen spritzen), liegt ungefähr bei 1 in 10.000. Wenn ein solcher Mann positiv testet, wie hoch ist die Wahrscheinlichkeit, dass er wirklich den Virus hat?

Aus den genannten Zahlen kann man erwarten, dass von je 10.000 Männern einer den Virus hat und mit an Sicherheit grenzender Wahrscheinlichkeit positiv testen wird. Von den 9.999, welche nicht infiziert sind, wird jedoch ebenfalls einer positiv testen (da die Falsch-positiv-Rate bei 0,01 % liegt). Also kann man zwei positive Testergebnisse erwarten: Eine Person davon hat den Virus, die andere nicht. (Diese Zahlen sind ungefähre Schätzungen, aber das ist nicht der Punkt hier.) In der Literatur sind nun eine Anzahl von Fällen beschrieben, in denen Personen, die keiner bekannten Risikogruppe angehörten, positiv testeten und in dem Glauben, dass sie mit absoluter Sicherheit HIV infiziert seien, auf einen langen Leidensweg gesandt wurden – vom

Verlust des Arbeitsplatzes zum ungeschützten Sexualkontakt mit anderen HIV-positiven Personen bis hin zum Selbstmord. Später stellte sich durch Zufall heraus, dass das Testergebnis ein falsch-positives war, das heißt, dass die Person gar nicht infiziert war (Stine, 1999). In einigen Fällen verklagten die Betroffenen die beratenden Ärzte, da diese die Möglichkeit von falsch-positiven Ergebnissen nicht erwähnt oder sogar bestritten hatten und das Testergebnis als absolut sicher erklärten (Gigerenzer, 2002).

Sind diese Fälle seltene, tragische Ausnahmen oder aber erzeugen AIDS-Berater regelmäßig diese Illusion der Gewissheit? Um diese Frage zu beantworten, suchte einer meiner Mitarbeiter als Klient 20 Beratungsstellen an deutschen Gesundheitsämtern in 20 verschiedenen deutschen Städten auf, um 20 HIV-Tests zu nehmen (ausführlich in Gigerenzer u.a., 1998). Dies war eine verdeckte Studie – die Berater wussten nicht, dass sie untersucht wurden. Vierzehn der 20 Berater waren Ärzte, die anderen Sozialarbeiter. Die obligatorische Beratung vor dem Test hat unter anderem zum Ziel, dem Klienten die Bedeutung eines positiven und negativen Testergebnisses zu erklären, hier stellte mein Mitarbeiter als Klient den Beratern jene Fragen, die jeder Klient stellen sollte. Beispielsweise: Kann es vorkommen, dass ich positiv teste, ohne den Virus zu haben? Wenn ja, wie oft passiert dies? Falls ein Mann wie ich, der in keiner Risikogruppe ist, positiv testet, wie hoch ist die Chance, dass ich wirklich den Virus habe?

Welche Information gaben die Berater dem Klienten über falsch-positive Ergebnisse? Sechzehn der 20 Berater versicherten, dass es keine falsch-positiven Ergebnisse gäbe. Auf Nachfragen, ob das wirklich nie passieren könnte, gaben drei zu, dass dies doch gelegentlich passiert, die anderen 13 beharrten aber darauf, dass falsch-positive Ergebnisse absolut unmöglich seien. Darüber hinaus erklärten auch fünf Berater, dass falsch-negative Ergebnisse nicht möglich seien.

Welche Information erhielt der Klient über den positiven Vorhersagewert des Tests, das heißt über die Wahrscheinlichkeit, dass er den Virus hat, wenn er positiv testet? (Erinnern Sie sich, es geht hier um einen Klienten, der kein Risikoverhalten praktiziert, und das Ergebnis „positiv“ bezieht sich auf wiederholte ELISA-Tests und einen Western Blot Test auf einer Blutprobe.) Zehn der 20 Berater versicherten dem Klienten, dass ein positives Ergebnis bedeutet, dass er mit absoluter Sicherheit den Virus hätte. Auch bei Nachfragen blieben sie dabei. Fünf weitere sagten ihm, dass es nicht absolut sicher sei, aber die Wahrscheinlichkeit sei bei 99,9 Prozent oder höher, und einer meinte 99 Prozent. Zwei Berater gaben ihm keine Antwort und wichen auch bei Nachfragen aus. Die restlichen zwei waren vorsichtiger und antworteten 90 Prozent oder höher. Einer von diesen beiden hatte allerdings auch zunächst versichert, dass nach einem positiven Test eine Infektion „vollständig sicher“ sei; erst auf eine Nachfrage wurde er nachdenklich. Die zweite Beraterin war die einzige, die dem Klienten richtig erklärte, dass eine niedrige Prävalenz (keine Risikogruppe) und niedrige Falsch-positiv-Rate zu einem relativ hohen Anteil von falsch-positiven unter allen positiven Ergebnissen führt.

Diese Studie zeigt, dass die aus den USA bekannt gewordenen Fälle von AIDS-Beratern, die Klienten in einer Illusion von Gewissheit wiegen, in Deutschland keine Ausnahme zu sein scheinen. Mehr als die Hälfte der deutschen Berater behaupteten falsche Sicherheiten. In jenen Fällen, in denen die Berater den Klienten über die Fehlerquoten des Tests tatsächlich informierten, drückten sie sich immer in Form von bedingten Wahrscheinlichkeiten aus (statt in natürlichen Häufigkeiten, siehe unten), mit der Konsequenz, dass einige Berater aus richtigen Wahrscheinlichkeitsangaben falsche Schlüsse zogen (Gigerenzer u.a., 1998). Beispielsweise erklärte ein Berater, dass die Sensitivität und die Spezifität des Tests jeweils 99,9 Prozent betragen und versicherte dem Klienten dann fälschlicherweise, daraus würde folgen, dass die Wahrscheinlichkeit einer Infektion, wenn der Klient positiv testet, ebenfalls 99,9 Prozent sei. Verwirrtes Denken mit Wahrscheinlichkeiten ist das Thema des nächsten Teils.

Die Illusion von Gewissheit kann Patienten teuer zu stehen kommen. Sie ist nicht auf AIDS-Beratung begrenzt. Wir fanden sie in den Aussagen von Sachverständigen vor Gericht über die Sicherheit von DNA-Analysen genauso wie in Informationsbroschüren mancher Gesundheitsorganisationen (Gigerenzer, 2002; Koehler u.a., 1995). Eine Analyse der Broschüren australischer Gesundheitsorganisationen über Mammographie-Screening zeigte beispielsweise, dass in keiner einzigen (!) der Broschüren über die Falsch-positiv-Rate berichtet wurde und ebenfalls nie über den positiven Vorhersagewert (Slaytor & Ward, 1998). Mit einer solchen Informationspolitik arbeitet man gezielt gegen das Ideal von „shared decision making“. Patienten, Ärzte und Gesundheitsbehörden müssen lernen, mit Unsicherheiten umzugehen statt diese zu verleugnen.

Wie kommuniziert man Risiken, sodass Arzt und Patient diese verstehen?

Effektive Risikokommunikation setzt voraus, dass Arzt und Patient die Illusion von Gewissheit hinter sich gelassen haben und versuchen, die Unsicherheiten zu verstehen und mit ihnen zu leben. Statistisches Denken ist hinreichend, um Unsicherheiten zu verstehen, nicht aber, um sie verständlich zu kommunizieren. Hierzu muss man wissen, welche Form der Information das menschliche Gehirn schnell versteht und welche es verwirrt. Die folgenden drei Formen gehören zu jenen, die regelmäßig Missverständnisse und Verwirrung herbeiführen: *Einzelfall-Wahrscheinlichkeiten*, *bedingte Wahrscheinlichkeiten* und *relative Risiken*. Überraschenderweise sind diese Formen zugleich jene, die in medizinischen Lehrbüchern, Broschüren und den Medien heute am meisten verwendet werden.

Einzelfall-Wahrscheinlichkeiten

Wenn Sie im Wetterbericht hören, dass es morgen mit einer Wahrscheinlichkeit von 30 Prozent regnet, dann hat man Ihnen eine Einzelfall-Wahrscheinlichkeit mitgeteilt. Die Wahrscheinlichkeit bezieht sich auf einen singulären Fall: den morgigen Tag, an dem es entweder regnet oder nicht. Wir glauben zu wissen, was diese Aussage bedeutet. Tatsächlich aber verstehen die Zuhörer mindestens drei verschiedene Dinge: Die einen meinen, dass es morgen in 30 Prozent der Zeit regnet, die anderen, dass es morgen in 30 Prozent der Gegend regnet, und schließlich wiederum andere, dass es an 30 Prozent der Tage wie morgen regnet. Nur Letzteres ist die Interpretation, welche die Meteorologen unterstellen. Diese Missverständnisse sind durch die Form der Kommunikation vorprogrammiert: Eine Einzelfall-Wahrscheinlichkeit gibt per Definition keine Referenzklasse an, auf die sich die Wahrscheinlichkeit bezieht. Das menschliche Gehirn des Zuhörers tendiert dagegen dazu, automatisch an konkrete Fälle zu denken, und eine bestimmte Referenzklasse anzunehmen, wie eine Klasse von Zeitpunkten, Orten oder Tagen. Sehen wir uns nun einen Fall aus der klinischen Praxis an, der illustriert, wie dort auf die gleiche Weise Missverständnisse entstehen, ohne dass sie im Allgemeinen bemerkt werden (Gigerenzer, 2001).

Ein Psychiater an der Universität von Michigan, USA, verschreibt depressiven Patienten regelmäßig das Arzneimittel Prozac. Dieses Medikament hat Nebenwirkungen im sexuellen Verhalten, wie Impotenz oder mangelnde Libido. Der Psychiater informierte die betreffenden Patienten regelmäßig darüber, dass „mit einer Wahrscheinlichkeit von 30 bis 50 Prozent“ eine Nebenwirkung auftreten wird. Als sie das hörten, waren viele Patienten doch beunruhigt, stellten aber keine weiteren Fragen, was den Psychiater schon immer verwundert hatte. Nachdem er über unsere Forschung am Max-Planck-Institut für Bildungsforschung gelesen hatte, begann er, auf andere Weise

über die Risiken zu informieren. Nun sagte er den Patienten, dass „bei drei bis fünf von zehn Personen“, die Prozac einnehmen, Nebenwirkungen auftreten. Mathematisch gesehen entsprechen diese Zahlen genau den Prozentsätzen, die er zuvor mitgeteilt hatte, psychologisch aber wirken sie ganz anders. Die Patienten, denen er die *Häufigkeiten* der unerwünschten Nebenwirkungen anstatt der *Prozentsätze* genannt hatte, waren deutlich weniger beunruhigt. Erst da wurde dem Psychiater klar, dass er nie nachgefragt hatte, wie seine Patienten eine „Wahrscheinlichkeit von 30 bis 50 Prozent“ verstanden hatten. Wie er schließlich herausfand, hatten viele gedacht, in 30 bis 50 Prozent ihrer sexuellen Aktivitäten würden sich Störungen einstellen. Jahrelang hatte der Psychiater nicht bemerkt, dass das, was er eigentlich sagen wollte, nicht das war, was seine Patienten verstanden.

Das zu Grunde liegende Problem ist dasselbe wie bei der Wettervorhersage. Bei der Aussage des Psychiaters, dass „mit einer Wahrscheinlichkeit von 30 bis 50 Prozent Nebenwirkungen auftreten können“, bleibt die Referenzklasse unklar: Bezieht sich der Prozentsatz auf eine Menge von Menschen (nämlich der Patienten, die Prozac einnehmen) oder auf eine Menge von Ereignissen (der sexuellen Aktivitäten einer bestimmten Person) oder auf irgendeine andere Menge? Der Psychiater nahm an, dass sich die Aussage auf seine Patienten bezog, während seine Patienten glaubten, sie beziehe sich auf ihre sexuellen Aktivitäten. Beide Seiten unterstellten eine andere Bezugsgruppe, jede aus ihrer Sicht.

Wie kann man dieses Problem bei der Kommunikation von Risiken vermeiden? Indem man grundsätzlich Aussagen in der Form von Häufigkeiten statt von Einzelfall-Wahrscheinlichkeiten trifft. Beispielsweise wird bei der Angabe von Häufigkeiten – „drei von zehn Patienten“ – die Bezugsmenge sofort klar, und die Gefahr von Missverständnissen ist deutlich geringer.

Das Problem der vieldeutigen Einzelfall-Wahrscheinlichkeiten ist nicht auf die Arzt-Patient-Interaktion beschränkt. Es betrifft ebenso die Interpretation von klinischen Studien. Beispielsweise wurden in einem MacArthur-Projekt die Vorhersage von gewalttätigem Verhalten durch hunderte von amerikanischen Psychiatern untersucht (Slovic & Monahan, 1995). Die Grundlage für die Vorhersagen waren detaillierte Entlassungsberichte. In den ersten Studien fragte man die Psychiater ausschließlich nach Einzelfall-Wahrscheinlichkeiten, wie „Was ist die Wahrscheinlichkeit, dass Mr. Smith in den ersten sechs Monaten nach der Entlassung gewalttätig wird?“ Als Reaktion auf unsere Forschung haben die Autoren in der neuesten Studie dann auch nach Häufigkeiten gefragt: „Wie viele von 100 Männern, deren relevante Merkmale denen von Mr. Smith gleichen, werden in den ersten sechs Monaten nach der Entlassung gewalttätig werden?“ Die Antworten auf die Häufigkeitsfrage waren systematisch niedriger. Wenn beispielsweise die mittlere Prognose bei der Einzelfallfrage 30 Prozent war, dann war sie bei der Häufigkeitsfrage nur 20 von 100 (Slovic, Monahan, & MacGregor, 2000).

Wie bei der Wettervorhersage und der Beratung über Nebenwirkungen ist die Frage nach Einzelfall-Wahrscheinlichkeiten auch hier vieldeutig. Die Wahrscheinlichkeit, dass ein Patient gewalttätig wird, kann wiederum auf mehrere Arten interpretiert werden: „Wenn man dem Patienten an 100 Wochenenden Freiheit gewährt, wie oft wird er eine Gewalttat begehen?“ oder aber „Wenn 100 Patienten wie ihm an einem Wochenende Freiheit gewährt wird, wie viele von ihnen werden eine Gewalttat begehen?“ Weitere Interpretationen sind denkbar. Mit einer Häufigkeitsfrage dagegen gibt man eine Bezugsmenge vor und schafft Klarheit an Stelle von Vieldeutigkeit.

Auch in Gerichtsprozessen werden folgenschwere Missverständnisse durch Sachverständige verursacht, welche in Einzelfall-Wahrscheinlichkeiten aussagen. Stellen Sie sich vor, Sie seien des Mordes angeklagt und stünden nun vor Gericht. Es gibt nur ein einziges Beweisstück gegen Sie – aber das hat es in sich: Ihr DNA-Profil stimmt mit dem einer Spur überein, die am Opfer gefunden wurde. Was bedeutet diese Übereinstimmung? Das Gericht beruft einen Sachverständigen, und dieser sagt Folgendes aus:

„Die Wahrscheinlichkeit, dass diese Übereinstimmung zufällig zu Stande kam, liegt bei 1 zu 100.000.“

Sie sehen sich schon hinter Gittern. Der Sachverständige könnte aber den gleichen Sachverhalt anders ausdrücken:

„Unter je hunderttausend Menschen wird sich bei einem eine Übereinstimmung zeigen.“

Mathematisch gesehen ist diese Aussage dieselbe wie die vorherige Wahrscheinlichkeitsaussage, aber psychologisch gesehen legt die Häufigkeitsform einem nahe zu fragen, wie viele Menschen es wohl gibt, die als Mörder in Frage kämen. Wenn Sie in einer Großstadt mit 1 Million erwachsenen Einwohnern leben, dürfte es demnach etwa zehn Einwohner geben, deren DNA-Profil mit dem der Spur am Opfer übereinstimmt. Diese Tatsache allein würde Sie kaum ins Gefängnis bringen.

Wie vor Gericht geht es auch in der Medizin nicht allein darum, die richtige Information zu kommunizieren, sondern auch in einer Form, die sie leicht verstehbar macht und Verwirrung vermeidet. Richtige Information falsch kommuniziert kann denselben Effekt haben wie falsche Information richtig kommuniziert. Einige Richter in England und den USA sind in jüngsten Jahren dieser Einsicht gefolgt und haben Sachverständige angewiesen, in Form von Häufigkeiten und nicht von Einzelfall-Wahrscheinlichkeiten auszusagen. In der Medizin sind mir bisher noch keine Richtlinien für effektive Risikokommunikation bekannt geworden.

Bedingte Wahrscheinlichkeiten

Eine bedingte Wahrscheinlichkeit gibt die Wahrscheinlichkeit eines Ereignisses A an, gegeben ein Ereignis B. Risiken und Unsicherheiten werden in der medizinischen Literatur wie auch in Informationsbroschüren häufig in der Form von bedingten Wahrscheinlichkeiten (oder Prozenten) dargestellt. Die Forschung zum statistischen Denken zeigt klar, dass die meisten Laien und manche Experten aus bedingten Wahrscheinlichkeiten – wie etwa die Sensitivität und die Falsch-positiv-Rate – falsche Schlüsse ziehen (Gigerenzer & Hoffrage, 1995). Betrachten Sie einmal das Brustkrebs-Screening durch Mammographie: Die Wahrscheinlichkeit, dass eine 50-jährige Frau Brustkrebs hat, beträgt nach einschlägigen Untersuchungen (Kerlikowske u.a., 1996) etwa 0,8 Prozent; die Wahrscheinlichkeit, dass das Mammogramm positiv ist wenn sie Brustkrebs hat, beträgt etwa 90 Prozent (Sensitivität); und die Wahrscheinlichkeit, dass das Mammogramm positiv ist wenn sie keinen Brustkrebs hat, ist etwa 7 Prozent (Falsch-positiv-Rate). Welche Wahrscheinlichkeit ergibt sich daraus dafür, dass eine Frau mit einem positiven Mammogramm wirklich Brustkrebs hat (positiver Vorhersagewert)?

Viele Menschen sind durch diese bedingten Wahrscheinlichkeiten verwirrt und können keine Antwort auf diese Frage finden. Wir können die Antwort jedoch leicht sehen, wenn wir eine andere Form der Kommunikation, an die das menschliche Gehirn während seiner langen Evolution angepasst ist: natürliche Häufigkeiten. Ich stelle nun die obigen Wahrscheinlichkeiten in natürlichen Häufigkeiten dar:

Von jeweils 1.000 Frauen haben 8 Brustkrebs. Von diesen 8 Frauen mit Brustkrebs werden 7 ein positives Mammogramm haben. Von den übrigen 992 Frauen, die keinen Brustkrebs haben, werden rund 70 dennoch ein positives Mammogramm haben. Betrachten wir nun jene Frauen, deren Mammogramm beim Screening positiv ausfiel. Wie viele von ihnen haben wirklich Brustkrebs?

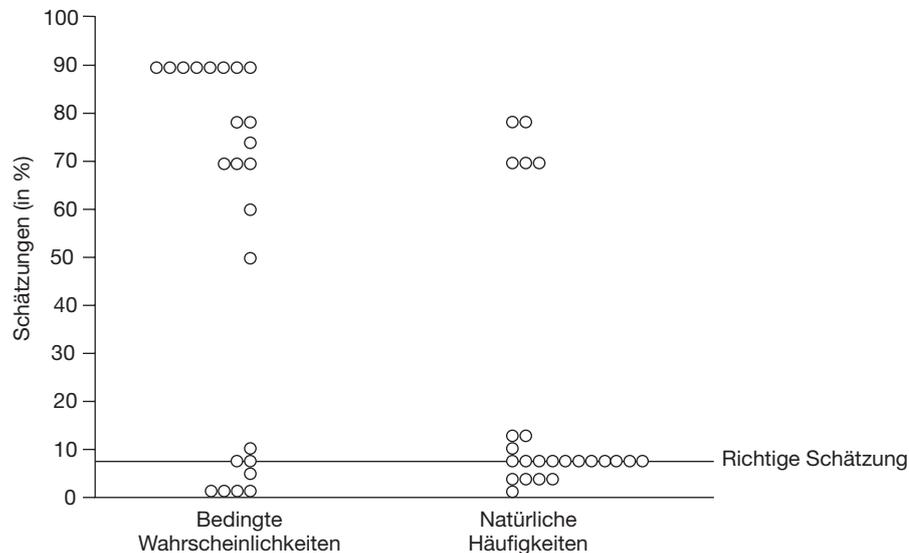


Abbildung 1: Bedingte Wahrscheinlichkeiten verwirren das Denken von Ärzten, natürliche Häufigkeiten können Einsicht erzeugen. 48 Ärzte beurteilten die Chance, dass Frauen mit einem positiven Screening-Mammogramm wirklich Brustkrebs hatten. Einer Hälfte der Ärzte (links) wurden die relevanten Informationen in Form von bedingten Wahrscheinlichkeiten (wie oft in medizinischen Lehrbüchern und Veröffentlichungen), der anderen Hälfte in natürlichen Häufigkeiten gegeben (rechts). Jeder Punkt repräsentiert die Einschätzung eines Arztes (nach Gigerenzer, 2002; Hoffrage & Gigerenzer, 1998).

Die hier mitgeteilten Informationen sind (abgesehen vom Ab- oder Aufrunden) genau die gleichen wie zuvor. Nun aber ist die Antwort viel leichter zu erkennen. Nur 7 von den insgesamt 77 Frauen mit positivem Mammogramm haben Brustkrebs; das entspricht 1 von 11 oder 9 Prozent.

Bedingte Wahrscheinlichkeiten machen es Ärzten schwer, die richtigen Schlussfolgerungen zu ziehen. Diesen Sachverhalt veranschaulicht eine Untersuchung, an der 48 deutsche Ärzte mit einer mittleren Berufserfahrung von 14 Jahren teilnahmen (Gigerenzer, 1996; Hoffrage & Gigerenzer, 1998). Achtzehn davon arbeiteten in Universitätskliniken, 16 in privaten oder öffentlichen Krankenhäusern und 14 hatten ihre eigene Praxis. Der einen Hälfte der Ärzte gaben wir die relevante Information über Mammographie-Screening in bedingten Wahrscheinlichkeiten und fragten sie, was die Wahrscheinlichkeit ist, dass eine Frau Brustkrebs wirklich hat, wenn sie positiv testet. Die anderen erhielten die Informationen in Form von relativen Häufigkeiten.

Wurden den Ärzten Wahrscheinlichkeiten genannt, gab es eine erschreckend geringe Übereinstimmung, wie der linke Teil von Abbildung 1 zeigt. Die Schätzungen schwankten zwischen 1 und 90 Prozent (!). Ein Drittel der Ärzte (8 von 24) schätzten die Wahrscheinlichkeit, dass eine Frau mit positivem Mammogramm tatsächlich Brustkrebs hat, mit 90 Prozent ein. Ein weiteres Drittel schätzte die Wahrscheinlichkeit auf 50 bis 80 Prozent. Der Rest (also ebenfalls ein Drittel) nannte eine Wahrscheinlichkeit von 10 Prozent oder weniger. Jede Patientin wäre mit Recht äußerst beunruhigt, wenn sie von dieser Diskrepanz erfahren würde. Nur zwei der 24 Ärzte

kamen auf den richtigen Wert, dem man mit der Regel von Bayes berechnen kann. Zwei weitere lagen mit ihren Schätzungen in der Nähe dieses Werts, aber nur zufällig, denn ihre Begründung war falsch. Beispielsweise verwechselte ein Teilnehmer die Falsch-positiv-Rate mit dem positiven Vorhersagewert.

Wie erging es den Ärzten mit natürlichen Häufigkeiten? Der rechte Teil von Abbildung 1 zeigt, dass hier keine derart beunruhigende Streuung der Ergebnisse vorliegt. Die Mehrzahl der Ärzte in dieser Gruppe lag mit ihrer Antwort richtig oder annähernd richtig. Nur fünf der 24 Ärzte gaben unrealistische Schätzungen von über 50 Prozent. Somit ließ sich allein durch die Form der Information bei vielen Ärzten, wenn auch nicht bei allen, Zahlenblindheit in Einsicht verwandeln.

In einer Reihe von Untersuchungen haben wir den Vorteil natürlicher Häufigkeiten bei Medizinstudenten wie Ärzten und bei Jurastudenten wie Richtern nachgewiesen (Gigerenzer, 2002; Hoffrage u.a., 2000). Die Verwendung von natürlichen Häufigkeiten an Stelle bedingter Wahrscheinlichkeiten ist eine einfache, schnelle und wirksame Methode, diagnostische Einsichten zu fördern.

Relative Risiken

Der Nutzen einer Behandlung wird oft in Form der relativen Risikoreduktion dargestellt. Beispielsweise wird mitgeteilt, dass Mammographie das Risiko, an Brustkrebs zu sterben, um 25 Prozent reduziert, oder dass prophylaktische Mastektomie das Risiko um 90 Prozent reduziert. Studien haben gezeigt, dass Patienten relative Risiken systematisch missverstehen. Viele nehmen fälschlicherweise an, eine 25-prozentige Reduktion würde bedeuten, dass von 100 Personen, die an der Behandlung teilnehmen, das Leben von 25 gerettet wird. Wie kann man die Information über den Nutzen so kommunizieren, dass eine Frau ihn leicht verstehen kann? Tabelle 1 zeigt die tatsächliche Mortalitätsreduktion durch Mammographie-Screening, ermittelt aus vier schwedischen randomisierten Studien mit rund 280.000 Frauen. Sie zeigt auch drei Formen, diesen Nutzen zu präsentieren.

Tabelle 1:
Die Reduktion der Mortalität durch Brustkrebs bei Frauen (über 40)
welche zehn Jahre lang am Mammographie-Screening teilnahmen

Nutzen des Mammographie-Screenings	
Kein Mammographie-Screening	4 Todesfälle (bei jeweils 1.000 Frauen)
Mammographie-Screening	3 Todesfälle (bei jeweils 1.000 Frauen)
<i>Drei Formen der Präsentation des Nutzens</i>	
Relative Risikoreduktion	Das Mammographie-Screening verringert das Risiko, an Brustkrebs zu sterben, um 25 Prozent.
Absolute Risikoreduktion	Das Mammographie-Screening verringert die Anzahl der Frauen, die an Brustkrebs sterben, um 1 pro 1.000, also um 0,1 Prozent.
Anzahl der notwendigen Behandlungen	Die Anzahl der Frauen, die zehn Jahre lang am Screening teilnehmen müssen, damit ein Todesfall verhindert wird, beträgt 1.000.

Die Ergebnisse stammen aus vier schwedischen randomisierten Studien mit 280.000 Frauen (die Werte in der Tabelle sind gerundet; nach Nyström u.a., 1996; Mühlhauser & Höldke, 1999).

Alle drei Darstellungen sind korrekt, suggerieren aber unterschiedlich hohen Nutzen und können bei unbefangenen Lesern verschiedene Emotionen und Bereitschaft zur Teilnahme am Screening auslösen. Die absolute Risikoreduktion ist der Anteil der Frauen, die ohne Screening starben, abzüglich des Anteils derer, die mit Screening starben. Diese absolute Verminderung beträgt 4 minus 3, also 1 von 1.000. Die relative Risikoreduktion ist die absolute Risikoreduktion, dividiert durch den Anteil der Frauen, die ohne Screening starben. Diese relative Risikoreduktion beträgt 1/4, also 25 Prozent. Die Anzahl der notwendigen Behandlungen, um einen Todesfall zu vermeiden („number needed to treat“), beträgt 1.000, denn einer von 1.000 Todesfällen wird durch das Screening verhindert.

Wenn die relative Risikoreduktion angegeben wird, sind Missverständnisse die Regel. Relative Werte enthalten keine Informationen über den absoluten Nutzen eines Verfahrens: Beispielsweise bedeutet eine 25-prozentige Verminderung, dass viele Menschenleben gerettet werden, wenn die betreffende Krankheit häufig ist, während nur wenige gerettet werden, wenn die Krankheit selten ist.

Eine transparente Information über den Nutzen einer Behandlung kann durch die Angabe absoluter Risiken oder der Anzahl notwendiger Behandlungen gewährleistet werden. Diese Formen entsprechen der Natur des menschlichen Gehirns, das dazu tendiert, in konkreten Personen zu denken, nicht in relativen Werten. Dennoch werden auch heute noch der Öffentlichkeit fast ausschließlich relative Risiken mitgeteilt und damit eine Überschätzung des Nutzens von Behandlungen bewusst in Kauf genommen. In der bereits erwähnten Analyse der Informationsbroschüren von Gesundheitsbehörden in Australien wurde der Nutzen immer nur in relativen Risiken mitgeteilt. Absolute Werte oder die Anzahl der notwendigen Behandlungen wurden nie mitgeteilt. In meiner Forschungsgruppe ist eben eine Analyse von 27 Mammographie-Broschüren deutscher Organisationen fertiggestellt worden. In Deutschland war die Qualität der Information etwas besser, aber nur geringfügig: Es gab eine einzige Broschüre, in der die Anzahl notwendiger Behandlungen angegeben war, und zwei, welche die absolute Risikoreduktion mitteilten (Kurzenhäuser, 2002).

Relative Werte entsprechen größeren Zahlen als absolute Werte und wirken dadurch eindrucksvoller. Dies ist einer der Gründe, warum sie bevorzugt werden. In einem Informationsblatt über hormonhaltige Wechseljahr-Medikamente, das in deutschen gynäkologischen Praxen auslag, wurde beispielsweise der potenzielle Vorteil der Hormone (eine Reduktion des Auftretens von Darmkrebs) in Form von relativen Risiken angegeben, der potenzielle Nachteil (Erhöhung des Auftretens von Brustkrebs) aber in der Form absoluter Risiken. Dies lässt Vorteile größer und Nachteile kleiner erscheinen. Das bewusste Erzeugen von systematischen Missverständnissen ist wohl die Ausnahme. Aber man kann an diesem Fall sehen, wie wesentlich die klare Kommunikation von Risiken für die Glaubwürdigkeit des Arztes und das Vertrauen unserer Gesellschaft in das medizinische System ist.

Shared Decision Making: Der Traum vom mündigen Patienten

Der Philosoph Immanuel Kant begann seinen im Jahre 1784 verfassten Aufsatz „Zur Beantwortung der Frage: Was ist Aufklärung?“ folgendermaßen: *Aufklärung ist der Ausgang des Menschen aus seiner selbstverschuldeten Unmündigkeit.*

Ich habe heute Abend über zwei Formen von Unmündigkeit gesprochen, die Illusion der Gewissheit und die mangelnde Fähigkeit, Risiken und Ungewissheiten zu verstehen und zu kommunizieren. Diese Formen der Unmündigkeit sind heute, mehr als zwei Jahrhunderte nach-

dem Kant diese Worte schrieb, immer noch verbreitet. Aber auch Ärzte haben ein elementares Bedürfnis für Aufklärung und einen Anspruch darauf, zu wissen, wie man Informationen kommuniziert, sodass der Patient versteht. Dies führt zu zwei Konsequenzen meines Vortrags:

- (1) *Richtlinien für die Ausbildung von Medizinstudenten.* Die Ausbildung der Medizinstudenten sollte ein Training in effektiver Risikokommunikation einschließen und auf ineffektive Formen aufmerksam machen. Angehende Ärzte sollten lernen, sich in natürlichen Häufigkeiten, absoluten Risiken und „number needed to treat“ auszudrücken, statt mit bedingten Wahrscheinlichkeiten, Einzelfall-Wahrscheinlichkeiten und anderen Formen kollektive Verwirrung zu stiften. Ein solches Training benötigt wenig Zeit und sein Effekt ist, wie unsere Untersuchungen gezeigt haben, über die Zeit hin stabil (Sedlmeier & Gigerenzer, 2001). Wenn man dagegen Studenten statistisches Denken nur in Form von bedingten Wahrscheinlichkeiten oder der Regel von Bayes in der Ausbildung vermittelt, vergessen sie das Gelernte oft schneller als sie brauchten, um es auf die Klausur zu lernen.
- (2) *Richtlinien für Broschüren und Zeitschriften.* Die zweite Konsequenz ist die Einführung von Richtlinien für Risikokommunikation in Informationsbroschüren der Gesundheitsorganisationen und in medizinische Zeitschriften. Dabei sollen sich die Organisationen verpflichten, die relevante Information über Nutzen, Nebenwirkungen und andere Risiken in einer Form darzustellen, die für das menschliche Gehirn leicht verständlich ist.

Ich möchte zum Abschluss nochmals betonen, dass die Illusion von Gewissheit und die Probleme mit der Risikokommunikation nicht allein in der Medizin auftreten. Vielmehr treten diese bei Laien genauso auf wie bei Experten, und in der Kommunikation über die Risiken neuer Technologien genauso wie über die Bedeutung von DNA-Beweismitteln vor Gericht. Wie bereits erwähnt, haben die ersten Richter in den USA und in Großbritannien entsprechende Richtlinien eingeführt, welche die Sachverständigen anweisen, die Information über DNA-Evidenz in natürlichen Häufigkeiten statt in Wahrscheinlichkeiten oder Likelihood-Quotienten vorzutragen – eben um die bekannte Verwirrung zu unterbinden.

Wie kann die Medizin von der Kunst der Risikokommunikation profitieren? Der Arzt wird treffsicherer. Der Patient wird besser geführt. Die medizinische Praxis erreicht höhere Qualität und Glaubwürdigkeit. Die einzigen, welche verlieren sind jene, die Scharlatanerie verkaufen.

Kehren wir zum Abschluss nochmals zu Kant zurück. Kant hat den Gedanken der Aufklärung in nur zwei Worten ausgedrückt: *sapere aude*. Habe den Mut zu wissen. Wenn wir in der Medizin einen mündigen Patienten wollen, dann müssen wir ihm helfen, den Mut zu finden, Ungewissheiten und Risiken zu verstehen und mit ihnen zu leben. Diesen Mut wird er nur finden, wenn er die Risiken auch versteht. Das Wissen darüber, wie man Risiken effektiv kommuniziert, haben wir. Jetzt ist der Mut gefragt, dieses Wissen in der Medizin umzusetzen.

Literatur

- Gigerenzer, G. (1996). The psychology of good judgment: Frequency formats and simple algorithms. *Medical Decision Making, 16*, 273–280.
- Gigerenzer, G. (2001). Der unmündige Patient. *Kursbuch, 145*, 132–144.
- Gigerenzer, G. (2002). *Das Einmaleins der Skepsis: Über den richtigen Umgang mit Zahlen und Risiken*. Berlin: Berlin Verlag.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.
- Gigerenzer, G., Hoffrage, U., & Ebert, A. (1998). AIDS counseling for low-risk clients. *AIDS Care, 10*, 197–211.
- Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine, 73*, 538–540.

- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, *290*, 2261–2262.
- Kerlikowske, K., Grady, D., Barclay, J., Sickles, E. A., & Ernster, V. (1996). Effect of age, breast density, and family history on the sensitivity of first screening mammography. *Journal of the American Medical Association*, *276*, 33–38.
- Koehler, J. J., Chia, A., & Lindsey, S. (1995). The random match probability (RMP) in DNA evidence: Irrelevant and prejudicial? *Jurimetrics Journal*, *35*, 201–219.
- Kurzenhäuser, S. (2002). *Brustkrebs-Screening: Welche Information über Risiken und Nutzen vermitteln deutsche Gesundheitsbroschüren?* Manuskript, Max-Planck-Institut für Bildungsforschung, Berlin.
- Mühlhauser, I., & Höldke, B. (1999). Übersicht: Mammographie-Screening. Darstellung der wissenschaftlichen Evidenz-Grundlage zur Kommunikation mit der Frau. *Sonderbeilage Arzneitelegramm*, *10*, 101–108.
- Nyström, L., Larsson, L.-G., Wall, S., Rutqvist, L., Andersson, I., Bjurstam, N., Fagerberg, G., Frisell, J., & Tabár, L. (1996). An overview of the Swedish randomised mammography trials: Total mortality pattern and the representativity of the study cohorts. *Journal of Medical Screening*, *3*, 85–87.
- Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*, 380–400.
- Slaytor, E. K., & Ward, J. E. (1998). How risks of breast cancer and benefits of screening are communicated to women: Analysis of 58 pamphlets. *British Medical Journal*, *317*, 263–264.
- Slovic, P., & Monahan, J. (1995). Probability, danger, and coercion: A study of risk perception and decision making in mental health law. *Law and Human Behavior*, *19*, 49–65.
- Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, *24*, 271–296.
- Stine, G. J. (1999). *AIDS update 1999: An annual overview of acquired immune deficiency syndrome*. Upper Saddle River, NJ: Prentice-Hall.