

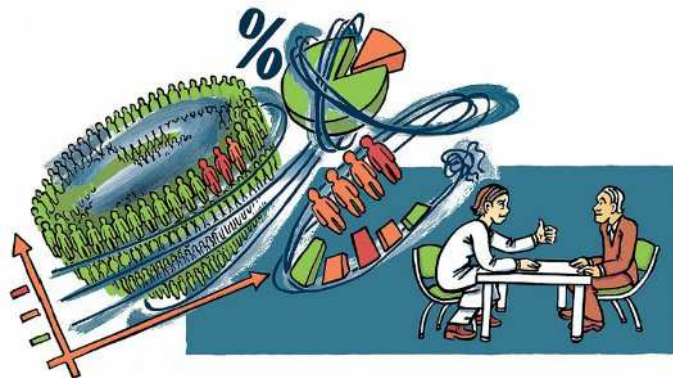
THEMEN DER ZEIT

Risikokommunikation: Risiken und Unsicherheiten richtig verstehen lernen

Dtsch Arztebl 2011; 108(9): A-448 / B-360 / C-360

Wegwarth, Odette; Gigerenzer, Gerd

Medizinische Fachzeitschriften und Broschüren verbreiten häufig intransparente Statistiken, die Ärzte irreführen, da diese in ihrer Ausbildung nicht darauf vorbereitet wurden.



Zeichnungen: Elke R. Steiner

Häufig geht die Forschung zur medizinischen Risikokommunikation davon aus, dass nur Patienten Schwierigkeiten mit dem Verständnis von Gesundheitsstatistiken haben. Aber auch Ärzte haben Probleme mit Zahlen. Viele lassen sich durch relative Risikoangaben verwirren, können den positiven Vorhersagewert (PPV) von Früherkennungstests nicht korrekt bestimmen und nehmen fälschlicherweise beim Bewerten von Früherkennungsmaßnahmen an, dass höhere Fünfjahresüberlebensraten eine geringe Sterblichkeit bedeuten. Diese Schwierigkeiten gibt es vor allem deshalb, weil Ärzte an Universitäten unzureichend im Umgang mit Risiken und Unsicherheiten ausgebildet werden und medizinische Zeitschriften und Broschüren statistische Informationen regelmäßig in missverständlicher Form verbreiten.

25 Prozent Risikoreduktion – was bedeutet das?

Anzeige

Den Nutzen einer Behandlung kann man in mindestens drei unterschiedlichen Formen ausdrücken: als relative Risikoreduktion (RRR), als absolute Risikoreduktion (ARR) oder als Anzahl der Personen, die behandelt werden müssen, um einen Todes-/Krankheitsfall zu verhindern (NNT = number needed to treat). Ein Beispiel soll das verdeutlichen: Im Jahr 1996 (3) wurden die ersten Zwischenergebnisse mehrerer randomisierter kontrollierter Studien zum Nutzen der Mammographie veröffentlicht. Die damaligen Ergebnisse zeigten, dass eine regelmäßige Teilnahme an der Mammographie über zehn Jahre das Risiko von Frauen (50 bis 69 Jahre), an Brustkrebs zu sterben, von etwa vier auf drei pro 1 000 Frauen reduzierte. In den drei unterschiedlichen Formen dargestellt, sieht das Ergebnis folgendermaßen aus:

- **RRR:** Eine regelmäßige Teilnahme an der Mammographie über zehn Jahre reduziert das Risiko, an Brustkrebs zu sterben, um 25 Prozent (= von vier auf drei).
- **ARR:** Eine regelmäßige Teilnahme an der Mammographie über zehn Jahre reduziert das Risiko, an Brustkrebs zu sterben, von etwa vier auf drei pro 1 000 Frauen, also um 0,1 Prozent (eine Frau pro 1 000).
- **NNT:** Wenn 1 000 Frauen zehn Jahre lang regelmäßig an der Mammographie teilnehmen, wird eine Frau weniger an Brustkrebs sterben.

Relative Risiken sind oft große Zahlen, absolute Risiken und Angaben zur NNT dagegen kleine Zahlen. Auch liefern relative Risiken im Gegensatz zu den anderen beiden Formen keinerlei Informationen zur Basisrate des Ereignisses mit. Dadurch verwirren sie viele Menschen und führen zur Überschätzung des Nutzens.

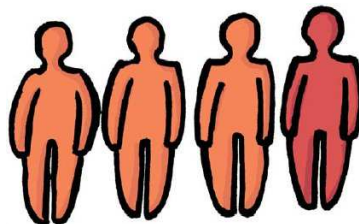
Der Chefarzt eines Luzerner Kantonsspitals befragte seine 15 Gynäkologen nach einem Interview mit Gerd Gigerenzer, was die 25-prozentige Risikoreduzierung durch Mammographie bedeute. Auf die Frage, wie viele Frauen von 1 000 weniger an Brustkrebs sterben, war ein Arzt der Meinung, 25 Prozent bedeute 2,5 von 1 000 Frauen, ein anderer hielt 25 von 1 000 Frauen für die richtige Antwort. Die gesamte Spanne der Antworten umfasste eine bis 750 Frauen von 1 000 (4).

Dieses Verständnisproblem gibt es nicht nur in der Schweiz. Zu Beginn eines Fortbildungskurses über Risikokommunikation legten wir einer Gruppe von 150 Gynäkologen dieselbe Frage vor (5). Die Ärzte konnten in einer TED-Abfrage zwischen vier Antwortalternativen wählen (eine, 25, 100 oder 250 von 1 000 Frauen). Zwei Drittel der Gynäkologen gaben die richtige Schätzung ab (eine von 1 000). 16 Prozent meinten jedoch, die 25-prozentige Reduktion bedeute 25 von 1 000 Frauen, und 15 Prozent waren der Ansicht, dass es 250 von 1 000 Frauen seien.

Nicht nur Gynäkologen fällt es schwer, den Nutzen von Screening richtig zu verstehen. Nachdem der PSA-Test Ende der 1980er-Jahre eingeführt worden war, entwickelte sich dieser zu einer häufig angebotenen individuellen Gesundheitsleistung zur Früherkennung von Prostatakrebs. Seit März 2009 liegen nun die ersten Ergebnisse zweier randomisierter Studien aus Europa und den USA zum Nutzen dieser Früherkennung vor. Während sich in der amerikanischen Studie keine Reduktion der krankheitsspezifischen Sterblichkeit durch die Teilnahme an der Früherkennung zeigte (6), fand die europäische Studie eine Reduktion von 3,7 auf 3,0 an Prostatakrebs gestorbenen Männern pro 1 000 (1). Diese Reduktion entspricht der einst Schlagzeilen machenden RRR von 20 Prozent (2). Eine positive Auswirkung der Teilnahme an der Früherkennung auf die Gesamtsterblichkeit wurde in beiden Studien nicht gefunden. Jedoch legt die europäische Studie das Ausmaß der Überbehandlungen offen: In der Gruppe mit Früherkennung wurden etwa 30 pro 1 000 Männer überdiagnostiziert und überbehandelt (1). Der Entdecker des PSA, Richard Ablin, bezeichnet daher das routinemäßige PSA-Screening als „Katastrophe für das Gesundheitssystem“ („Süddeutsche Zeitung“ vom 12. März 2010). Aufgrund des kritischen Nutzen-Schaden-Verhältnisses empfehlen medizinische Leitlinien seit langem eine Aufklärung der Patienten über den Nutzen und Schaden der PSA-Tests (7). Die „Stiftung Warentest“ prüfte die Qualität der Beratung von Urologen.

Schwierigkeiten im Umgang mit Risikoangaben

In einer verdeckten Studie ließ sich ein 60-jähriger Mann (selbst Arzt) von 20 zufällig ausgesuchten Urologen in Berlin über den PSA-Test beraten (8). Von den 20 befragten Urologen konnten lediglich zwei



relevante Informationen zum Nutzen und Schaden vermitteln, vier weitere kannten

einen Teil der Informationen. Eine Mehrheit von 14 Urologen erfüllte die Mindestanforderungen der Leitlinien ihrer Fachgesellschaft nicht: Sie behaupteten fälschlicherweise, dass die lebensverlängernde Wirkung der PSA-Testung wissenschaftlich bestätigt sei, und waren sich keinerlei Nachteile bewusst. Gynäkologen und Urologen sind keine Ausnahmen. Schwierigkeiten im Umgang mit Risikoangaben wurden auch für andere medizinische Fachrichtungen und über Länder hinweg gefunden (9–12).

Sensitivität, Spezifität und positiver Vorhersagewert

Nach Erhalt eines positiven Testergebnisses fragen sich viele Patienten besorgt, wie wahrscheinlich es nun ist, dass sie tatsächlich erkrankt sind. Darüber Auskunft geben soll der positive Vorhersagewert (PPV = positive predictive value). Der PPV ergibt sich aus der Prävalenz einer Erkrankung sowie der Sensitivität und der Falschpositivrate (1-Spezifität) eines Tests. Um Patienten nicht unnötigen Ängsten, Überdiagnosen und Überbehandlungen auszusetzen, sollte ein Arzt wissen, wie man den PPV errechnet.

Haemocult-Test (FOBT = Fecal occult blood test): Hoffrage und Gigerenzer (13) testeten 48 Ärzte verschiedener Fachrichtungen mit einer durchschnittlichen Berufserfahrung von 14 Jahren. Ärzte aus Universitätskliniken und Chefärzte waren in dieser Stichprobe leicht überrepräsentiert. Die Ärzte erhielten folgende Informationen zur Darmkrebs-Früherkennung mit Hilfe des Haemocult-Tests in Form von bedingten

Wahrscheinlichkeiten: Sensitivität 50 Prozent, Falschpositivrate drei Prozent, Prävalenz 0,3 Prozent. Anschließend wurden die Ärzte gebeten, die Wahrscheinlichkeit von Darmkrebs nach Erhalt eines positiven Testbefunds einzuschätzen. Die Antworten variierten von ein bis 99 Prozent. Die Mehrheit der Ärzte gab eine Wahrscheinlichkeit von 50 Prozent (Sensitivität) an. Vier der 48 Ärzte zogen die Falschpositivrate von der Sensitivität ab und gelangten so zu einer Einschätzung von 47 Prozent. Die tatsächliche Wahrscheinlichkeit beträgt jedoch fünf Prozent.

HIV-Beratung: Auf einer Konferenz zum Thema Aids im Jahr 1987 berichtete der damalige Senator Floridas, Lawton Chiles, dass von 22 Blutspendern in Florida, die darüber informiert wurden, dass sie mit dem ELISA-Test positiv getestet worden waren, sieben Suizid begangen hätten. Ein medizinisches Schriftstück, das diese Tragödie Jahre später dokumentierte, informierte den Leser darüber, dass „selbst wenn die Ergebnisse des Aidstests (ELISA- und Western-Blotting-Test) positiv ausfallen, die Wahrscheinlichkeit einer Infektion nur 50 : 50 beträgt“ (14). Warum ist die Wahrscheinlichkeit bei einer Sensitivität von 99,9 Prozent und einer Spezifität von 99,99 Prozent so gering? Der Grund dafür liegt in der niedrigen Basisrate von Aids in der Population mit geringem Risikoverhalten (circa 0,01%). Wenn sich 10 000 Menschen mit niedrigem Risikoverhalten nun testen lassen, wird die eine tatsächlich infizierte Person aufgrund der hohen Sensitivität sehr wahrscheinlich positiv auf HIV getestet. Von den 9 999 nichtinfizierten Menschen wird aber aufgrund der Falschpositivrate von etwa 0,01 Prozent ebenfalls eine Person positiv getestet werden. Das ergibt zwei Personen mit einem positiven Test, von denen aber nur eine Person infiziert ist.

In einer verdeckten Studie in 20 öffentlichen Gesundheitszentren in Deutschland (15) wurde bei Aidsberatern eine weitgehende Unkenntnis zu diesem Sachverhalt festgestellt. Der dazu eingesetzte „Patient“ fragte unter anderem nach der Wahrscheinlichkeit einer positiven Testung, wenn man gar nicht infiziert sei. Obwohl der Patient klarmachte, dass er zur Gruppe mit geringem Risikoverhalten gehöre, behaupteten trotzdem 16 Berater, dass falsche Testergebnisse nie vorkämen; drei von ihnen nahmen ihre Aussage zurück, als der Patient nachfragte, ob sie sich absolut sicher seien. Lediglich drei Berater teilten dem Patienten mit, dass es falschpositive Ergebnisse geben könne, da die Spezifität nicht perfekt – wenngleich sehr hoch – sei.



Die Schwierigkeit, den PPV einzuschätzen, ist nicht auf Deutschland beschränkt. In einer australischen Studie gaben 13 von 50 Ärzten an, sie könnten den positiven Vorhersagewert beschreiben, bei der anschließenden direkten Befragung gelang dies aber tatsächlich nur einem Arzt (16). Vergleichbares wurde für die Mitglieder der amerikanischen Nationalen Akademie für Neuropsychologie berichtet (17). Ghosh und Ghosh (18) geben einen Überblick zu weiteren Studien, die zeigen, dass nur wenige Ärzte in der Lage sind, aus den relevanten

Statistiken den positiven Vorhersagewert abzuschätzen.

Fünfjahresüberlebensrate mit geringem Aussagewert

Rudy Giuliani, der ehemalige Bürgermeister von New York, stellte im Jahre 2007 während seines Wahlkampfs zum Präsidenten der USA die Vorteile des amerikanischen Gesundheitssystems gegenüber dem britischen System heraus: „Vor fünf, sechs Jahren hatte ich Prostatakrebs. . . . Meine Chance, Prostatakrebs in Amerika zu überleben? 82 Prozent! Meine Chance, Prostatakrebs in Großbritannien zu überleben? Gerade mal 44 Prozent – unter einem verstaatlichten Gesundheitssystem.“

Die hier von Giuliani genutzten Prozentwerte waren die damaligen Fünfjahresüberlebensraten. Häufig werden steigende Überlebensraten mit sinkenden Sterblichkeitsraten gleichgesetzt und dann als untrügliches Zeichen für den Erfolg von früherem Erkennen und Behandeln im Kampf gegen Krebs gewertet. Diese Annahme ist jedoch nicht korrekt. Der Zusammenhang zwischen den Fünfjahresüberlebensraten und den Sterblichkeitsraten für die 20 häufigsten Tumoren in den USA ist gleich null (19). Obwohl die Überlebensstatistik ein valides Maß für Vergleiche von Krebstherapien in randomisierten Studien ist, ist sie es nicht für Vergleiche von Gruppen, deren Erkrankung unterschiedlich diagnostiziert wurde (Früherkennung versus Symptome). Der Grund dafür liegt hauptsächlich in zwei systematischen Verzerrungen: dem „Vorlaufzeit-Bias“ und dem „Überdiagnose-Bias“. Das „Vorlaufzeit-Bias“ entsteht dadurch, dass die Früherkennung den Diagnosezeitpunkt lediglich vorverlegt und sich die Überlebenszeit so nur scheinbar verlängert. Weil durch die Früherkennung auch sehr langsam wachsende oder nichtprogressive Tumoren mit in die Überlebensstatistik einbezogen werden, kommt es überdies zum „Überdiagnose-Bias“. Durch die Vorverlegung des Diagnosezeitpunkts und den Einschluss von Tumoren mit günstiger Prognose wird die Fünfjahresüberlebensrate durch die Früherkennung künstlich aufgebläht, ohne dabei zwangsläufig eine reale Entsprechung zu einer reduzierten Sterblichkeit zu haben. Aus diesen Gründen eignet sich die Fünfjahresüberlebensrate nicht dazu, den Effekt von Früherkennung einzuschätzen, weshalb Institutionen wie das National Cancer Institute empfehlen, stattdessen die krankheitsspezifische Sterblichkeit zu nutzen (20).

Dieser Sachverhalt ist Ärzten oft nicht bekannt. Wir befragten 65 Urologen, Gynäkologen und Internisten zu ihrem Verständnis von Fünfjahresüberlebensraten (21). In

unterschiedlichen Szenarien gaben wir jedem Arzt Informationen zu Veränderungen der Fünfjahresüberlebensrate oder der krankheitsspezifischen Sterblichkeit. Die jeweilige Information bezog sich dabei immer auf den Effekt ein und derselben Früherkennung. Nach jedem Szenario fragten wir, ob sie ihrem Patienten die Früherkennung empfehlen würden. Wurden den Ärzten Fünfjahresüberlebensraten vorgelegt, gaben etwa zwei Drittel an, dass sie die Früherkennung empfehlen würden, und begründeten diese Entscheidung mit dem deutlichen Anstieg der Überlebensrate in der Gruppe mit der Früherkennung. Wurde den Ärzten die krankheitsspezifische Sterblichkeit vorgelegt, waren weniger als zehn Prozent der Befragten bereit, die entsprechende Früherkennung zu empfehlen. Nach „lead-time bias“ oder „overdiagnosis bias“ befragt, konnten weniger als zehn Prozent korrekt erklären, was diese Bias bedeuten.

Gründe für das mangelnde Risikoverständnis

Die Ursachen für das mangelnde Risikoverständnis liegen nicht beim Arzt allein. Eine jahrzehntelange vernachlässigte Ausbildung im Verständnis von Statistiken und Risikokommunikation im Medizinstudium und in der ärztlichen Weiterbildung haben entscheidend zu der hier beschriebenen Situation beigetragen. Auch medizinische Fachzeitschriften und Broschüren verbreiten intransparente Statistiken, die Ärzte irreführen, da sie in ihrer Ausbildung nicht darauf vorbereitet wurden. Nuovo, Melnikow und Chang (22) analysierten 359 Artikel, veröffentlicht in: *Annals of Internal Medicine*, *British Medical Journal*, *Journal of the American Medical Association*, *The Lancet* und *The New England Journal of Medicine*. Nur 25 Artikel stellten die Ergebnisse auf eine transparente Art und Weise dar und berichteten eine absolute Risikoreduktion oder die NNT. Das Institut für Wirtschaftlichkeit im Gesundheitswesen analysierte 175 an Ärzte gerichtete Informationsblätter der Pharmaindustrie (23). Das Ergebnis spricht für sich: Die Zusammenfassungen konnten in nur acht Prozent der Fälle bestätigt werden, Originalstudien wurden oft nicht genannt oder waren nicht auffindbar, und in 92 Prozent der Fälle waren Schlüsselergebnisse der Originalstudien systematisch verfälscht und wichtige Details unterschlagen. Ein Prospekt von Bayer gab beispielsweise an, dass die Wirkung des Potenzmittels Levitra (Vardenafil) bis zu fünf Stunden anhält – ohne zu erwähnen, dass diese Statistik auf Studien mit betäubten Hasen basiert.



*Dr. rer. nat. Odette Wegwarth
Prof. Dr. phil. Gerd Gigerenzer
Harding-Zentrum für Risikokompetenz
Max-Planck-Institut für Bildungsforschung, Berlin*

@Literatur im Internet:
www.aerzteblatt.de/it0911

KEIN BUCH MIT SIEBEN SIEGELN

Ärzte sollten Risiken und Unsicherheiten richtig verstehen und dem Patienten verständlich erklären können. In drei kurzen Folgeartikeln im Deutschen Ärzteblatt sollen einfach erlernbare Techniken vermittelt werden, die helfen sollen, die angesprochene Verwirrung durch Statistik in Einsicht zu verwandeln. Statistisches Denken und transparente Risikokommunikation sind kein Buch mit sieben Siegeln, sondern eine Kunst, die schnell und einfach erlernbar ist.

ALLERLEI WAHRSCHEINLICHKEITEN

- Als im März 2009 die ersten Ergebnisse der europäischen randomisierten Studie zur Prostatakrebs-Früherkennung publiziert wurden (1), bezifferte eine dazugehörige Pressemitteilung den Nutzen mit 20 Prozent (2). Bedeutet dies, dass von 100 Männern, die regelmäßig zur Früherkennung gehen, 20 weniger an Prostatakrebs sterben?
- Die HIV-Tests ELISA und Western-Blotting-Test haben zusammen ungefähr eine Sensitivität von 99,9 Prozent und eine Spezifität von 99,99. Wenn eine Person bei einer Routineuntersuchung positiv getestet wird, folgt dann daraus, dass sie mit 99,9-prozentiger Wahrscheinlichkeit infiziert ist?
- In den USA leben fünf Jahre nach der Diagnose „Prostatakrebs“ immer noch 98 Prozent der Männer, in Großbritannien jedoch nur 77 Prozent. Sollten sich britische Männer deshalb besser in den USA behandeln lassen?