

Gynäkologe 2018 · 51:370–379
<https://doi.org/10.1007/s00129-018-4199-3>
 Online publiziert: 12. Februar 2018
 © Der/die Autor(en) 2018. Dieser Artikel ist eine Open-Access-Publikation.

Redaktion
 M. Kiechle, München
 W. Jonat, Kiel
 A. Quante, München



O. Wegwarth

Max-Planck-Institut für Bildungsforschung, Center for Adaptive Rationality/Harding Center for Risk Literacy, Berlin, Deutschland

Brustkrebsfrüherkennung – Nutzen und Risiken richtig kommunizieren

Eine effiziente Gesundheitsversorgung braucht gut informierte Ärzte und Patienten. In unserem gegenwärtigen Gesundheitssystem gibt es beide nicht. Der Großteil von ÄrztInnen kennt nicht den Unterschied zwischen absoluten und relativen Risiken, ist außerstande, den positiven Vorhersagewert eines Tests zu berechnen, und weiß nicht, dass die Fünfjahresüberlebensrate im Kontext der Bewertung von Früherkennung eine invalide Statistik ist. Damit sind sie nicht in der Lage, Früherkennungen objektiv zu beurteilen und PatientInnen all jene Informationen zu vermitteln, die sie für eine informierte Entscheidung brauchten.

Patienten gelten als das Problem unserer modernen Gesundheitsversorgung: Sie sind ein unwissendes, ängstliches, unwilliges Volk mit ungesunden Lebensgewohnheiten. Sie verlangen Medikamente, die von Berühmtheiten im Fernsehen angepriesen werden, bestehen auf unnötige und/oder teure Tests und Behandlungen und werden am Ende noch zu Klägern. Im Angesicht explodierender Gesundheitskosten haben die uninformierten und damit kostenintensiven Patienten Aufmerksamkeit erweckt. Zum Beispiel: Bei fast 10 Mio. Frauen in den USA wurden unnötige Pap-Abstriche zur Zervixkarzinomerkenung vorgenommen – unnötig deshalb, weil diese Frauen nach Hysterektomie gar keine Zervix mehr hatten [1]. Unnötige Pap-Abstriche schaden der Patientin nicht, aber das Gesundheitssystem vergeudet damit finan-

zielle Ressourcen, die an anderen Stellen für bessere Gesundheitsversorgung hätten eingesetzt werden können. Schätzungsweise eine Million US-amerikanische Kinder werden jedes Jahr unnötigen CTs (Computertomographien) unterzogen [2]. Sie bedeuten nicht nur Geldverschwendung, sondern richten auch Schaden an: Die ungefähr 70 Mio. CT-Scans, die pro Jahr in den Vereinigten Staaten durchgeführt werden, haben aufgrund der hohen Strahlenbelastung schätzungsweise 29.000 Krebserkrankungen zur Folge [2]. Warum lassen Frauen völlig unnötige Untersuchungen über sich ergehen, warum schützen Eltern ihre Kinder nicht vor einer unnützen Strahlendosis? Vermutlich weil sie nicht über den Schaden informiert sind und dem landläufigen Irrglauben aufsitzen, dass mehr Untersuchungen auch mehr Gesundheitsgewinn bedeuten. Als eine Zufallsstichprobe von 500 AmerikanerInnen gefragt wurde, ob sie lieber 1000 Dollar in bar oder eine Ganzkörper-CT gratis hätten, entschieden sich 3 von 4 für die CT [3].

Uninformierte PatientInnen gibt es nicht nur in den USA. Eine repräsentative Studie [4] mit 10.228 Personen aus 9 europäischen Ländern ergab, dass 89 % der Männer und 92 % der Frauen den Nutzen von PSA(prostata-spezifisches Antigen)- und Mammographie-Tests um ein Zehnfaches, ein Hundertfaches oder noch mehr überschätzten oder diesen schlicht nicht wussten. Warum aber wissen PatientInnen, die sich heute wie nie zuvor über Gesundheitsfragen informieren, nicht Bescheid?

Antworten auf diese Fragen reichen von der Auffassung, dass PatientInnen

nicht intelligent genug seien, um medizinische Sachverhalte zu verstehen, bis hin zu der Meinung, sie wollten einfach keine Zahlen hören. Eine nicht zu unterschätzende Zahl von Gesundheitspsychologen ergänzt die Liste ferner mit Willensschwäche und dem Wunschdenken [5]. Folglich sind die empfohlenen Heilmittel paternalistisch: Dem unmündigen Patienten muss dabei „geholfen“ werden, den rechten Weg der Gesundheit zu betreten [6]. Ein deutlich ernster zu nehmendes Problem scheint aber die umfassende Missinformation von PatientInnen zu sein, die bei ÄrztInnen beginnt, sich über Patientenbroschüren fortsetzt und auch vor den Medien nicht haltmacht.

Kommunizieren von Nutzen und Schaden

Im Jahr 1996 wurden die Ergebnisse von 4 randomisierten Studien [7] zur Mammographie-Früherkennung, an denen etwa 280.000 Frauen teilgenommen hatten, analysiert. Unter anderem zeigte sich, dass von je 1000 Frauen, die über 10 Jahre am Screening teilgenommen hatten, 3 an Brustkrebs verstorben waren, und von je 1000 Frauen, die nicht am Screening teilgenommen hatten, 4 an Brustkrebs verstorben waren. Eine weitere Analyse [8, 9] zeigte vergleichbare Effekte: Pro 1000 Frauen bestand ein Unterschied von 5 an Brustkrebs verstorbenen Frauen ohne Früherkennung versus 4 an Brustkrebs verstorbenen Frauen mit Früherkennung. Ein in 2006 veröffentlichter Cochrane-Review, der diese und weitere randomisierte Studien mit nun insgesamt etwa 500.000 Frauen

zusammenfasste, kam zu einer noch geringeren absoluten Risikoreduktion: Von 2000 Frauen, die regelmäßig über 10 Jahre hinweg an der Mammographie teilnahmen, verstarb eine Frau (11 vs. 10 Frauen) weniger an Brustkrebs [9]. Erstmals wurde auch der Schaden der Mammographie beziffert: Für jede Frau, die durch die Teilnahme an der Früherkennung vor dem Brustkrebstod bewahrt wurde, wurden gleichzeitig 10 andere Frauen als Ergebnis ihrer Teilnahme an der Mammographie mit Brustkrebs überdiagnostiziert und überbehandelt. Sowohl die Nutzenbewertung als auch die Schadenbewertung der Mammographie wurden im aktuellsten Cochrane Review [10] – nun mit randomisierten, kontrollierten Studien, die rund 600.000 Frauen einschlossen – erneut bestätigt.

Was bedeuten Überdiagnose und Überbehandlung? Überdiagnose ist kein falsch-positiver Befund, sondern das Entdecken von Gewebeveränderungen, welche der pathologischen Definition eines Tumors entsprechen, sich jedoch nicht zu einem Tumor entwickeln, der Symptome oder gar den Tod verursacht. Die Konsequenz der Überdiagnose ist Überbehandlung – Operationen, Chemotherapien und Strahlenbehandlungen, welche der überdiagnostizierten Patientin keinerlei Überlebensnutzen, sondern nur körperlichen und seelischen Schaden bringen.

Werden Frauen, die an der Mammographie teilnehmen wollen, über

diese Fakten von ihren ÄrztInnen informiert? Um mehr darüber zu erfahren, führten wir im Jahr 2008 – fast 2 Jahre nach der Veröffentlichung des im Jahr 2006 erschienenen Cochrane-Reviews zur Mammographie [9] – eine Studie mit GynäkologInnen in verschiedenen Großstädten Deutschlands durch und befragten sie explizit zu Nutzen und Schaden dieser Früherkennung [11]. Von den 20 GynäkologInnen, die bereit waren, Auskunft zu erteilen, empfahlen 17 die Mammographie mit Nachdruck und betonten, dass sie eine sichere und wissenschaftlich fundierte Früherkennungsmethode sei. Nur 7 ÄrztInnen konnten Zahlen zum erfragten Nutzen nennen. Sie gaben an, dass durch die Mammographie die Brustkrebssterblichkeit um 20–50% gesenkt würde. Den potenziellen Schaden der Mammographie beschrieb die Mehrzahl der GynäkologInnen als „vernachlässigbar“. Nur 3 GynäkologInnen nannten Zahlen zu spezifischen Nachteilen, von diesen Angaben waren 2 falsch. Die Zahlen, die wir für den Schaden erhielten, waren im Gegensatz zu den Zahlen zum Nutzen jedoch keine Prozentzahlen, sondern lauteten zum Beispiel 1 Frau pro 26.000. Der Hauptschaden der Mammographie – Überdiagnose und Überbehandlung – wurde von keinem bzw. keiner der befragten GynäkologInnen erwähnt.

Die Ergebnisse der Studie deuten zweierlei an: Eine Frau, die sich über

den Nutzen und Schaden der Mammographie informieren möchte, würde zum einen entweder gar keine oder missverständliche Zahlen von ihrem Gynäkologen bzw. ihrer Gynäkologin erhalten. Und zum anderen würde sie, für den Fall, dass sie Zahlen erhielte, der Täuschung des „*mismatched framings*“ [2] aufsitzen. „*Mismatched framing*“ bedeutet, dass der Nutzen und Schaden von Behandlungen in unterschiedlichen „Währungen“ kommuniziert wird. Potenzielle „Währungen“, in denen der Nutzen und Schaden dargestellt werden kann, sind:

- das relative Risiko („*relative risk*“, RR),
- das absolute Risiko („*absolute risk*“, AR) oder
- die Anzahl notwendiger Behandlungen („*number needed to treat*“, NNT), die erfolgen müssen, um einen Krankheits- oder Todesfall zu verhindern (entspricht: 1/ARR).

Um ein relatives Risiko handelte es sich beispielsweise bei all jenen numerischen Angaben, die uns die GynäkologInnen zum Nutzen der Mammographie benannten. Die Schätzungen zur Senkung der Brustkrebssterblichkeit lagen hier bei 20–50%. Während es für die 50% keinerlei absolute Entsprechung in der gegenwärtigen Literatur zur Beurteilung der Mammographie für die Gesamtbevölkerung gibt, entsprechen die 20% der Analyse von Nyström et al. [8], in

Hier steht eine Anzeige.

der 5 von 1000 Frauen in der Gruppe ohne Mammographie und 4 von 1000 Frauen in der Gruppe mit Mammographie an Brustkrebs verstarben. Die 20% ergeben sich, indem die Referenzklasse (hier: 1000 Frauen) weggelassen und nur noch das Verhältnis der Reduktion von 5 (= 100%) auf 4 (= 80% von 5) betrachtet wird. Man spricht von einer *relativen Risikoreduktion (RRR)*. Das Problem relativer Risikoangaben ist jedoch, dass durch das Weglassen der Referenzklasse (hier: 1000 Frauen) und das Ignorieren der absoluten Größe der Basisrisiken (hier: 5 und 4) der absolute und damit letztlich der klinisch relevante Effekt nicht mehr abschätzbar ist. Denn 20% ergäben sich z. B. auch bei einer Reduktion von 500 auf 400 pro 1000, bei einer Reduktion von 0,0005 auf 0,0004 pro 1.000.000 und auch bei einer Reduktion von 200 auf 160 pro 147.839. Die tatsächliche absolute Risikoreduktion (ARR) der Brustkrebssterblichkeit von 0,1% (von 5/1000 auf 4/1000 = 1/1000) lässt sich aus den 20% also nicht schlussfolgern.

» Bei relativen Risikoangaben ist der klinisch relevante Effekt nicht mehr abzuschätzen

Wenn sich aus den relativen Risiken ohne Wissen um die absoluten Kennzahlen keine Beurteilung der klinischen Relevanz der Ergebnisse schließen lässt, warum werden sie dann dennoch in der Kommunikation verwendet? Der Grund liegt in der unterschiedlichen Wirkung von relativen und absoluten Risikoangaben. Im Gegensatz zu absoluten Risikoangaben produzieren relative Risikoangaben große Zahlen (RRR: 20% vs. ARR: 0,1%). Dies führt dazu, dass Laien und ÄrztInnen den so dargestellten Nutzen um das bis zu Hundertfache und mehr überschätzen [12]. In der Schweiz wurden 15 GynäkologInnen gefragt, was eine 25%ige Risikoreduktion im Hinblick auf den Nutzen des Mammographie-Screenings bedeute [13]. Ein Arzt dachte, dass dadurch 2,5 von 1000 Frauen weniger an Brustkrebs versterben, ein anderer nahm an, es seien 25 von 1000. Die Gesamtheit der Antworten reichte von 1–750 pro 1000 vor dem

Brustkrebstod geretteten Frauen. Bereits 1992 fanden Naylor et al. [14], dass Ärzte die Effektivität einer Behandlung dann unrealistisch hoch bewerteten, wenn der Nutzen in Form einer RRR beschrieben wurde, nicht aber, wenn dies in Form der ARR oder der NNT geschah. Eine Vielzahl weiterer Studien dokumentiert die täuschende Wirkung relativer Risikoangaben auf verschiedenste ExpertInnen im Gesundheitswesen [15, 16]. Die Fehlinterpretation von RRR aufseiten der ÄrztInnen bleibt nicht ohne Auswirkungen aufseiten der PatientInnen [12]. Aufgrund ihrer „vergrößernden“ Wirkung werden relative Risikoangaben besonders häufig zu Darstellungen des Nutzens medizinischer Maßnahmen verwendet [17]; sie machen ÄrztInnen und PatientInnen gleichermaßen glauben, dass ihnen die Teilnahme an einer Früherkennung oder die Einnahme eines Medikaments besonders viel bringe. Im Gegenzug wird für die Kommunikation des Schadens häufig das Format des absoluten Risikoanstiegs [17] verwendet, so auch in unserer oben dargestellten Studie, in der die GynäkologInnen alle 3 numerischen Angaben zum Schaden als absoluten Risikoanstieg (z. B. 1 von 26.000 Frauen) kommunizierten. Absolute Risikoangaben und auch die Anzahl notwendiger Behandlungen produzieren für gewöhnlich kleine Zahlen und lassen damit den Schaden als vernachlässigbar erscheinen. Wird der Nutzen in relativen Risiken und der Schaden in absoluten Risiken kommuniziert, dann nennt man dies „*mismatched framing*“ [18]. Eine Analyse von 359 Artikeln, veröffentlicht in hochrangigen medizinischen Periodika, wie *Annals of Internal Medicine*, *British Medical Journal (BMJ)*, *Journal of the American Medical Association (JAMA)*, *The Lancet* und *The New England Journal of Medicine*, zeigte im Jahr 2002, dass nur 25 Artikel ihre Ergebnisse auf transparente Art und Weise darstellten und eine absolute Risikoreduktion bzw. eine NNT angaben. Auch die GynäkologInnen unserer Studie [11] nutzten ausschließlich die Technik des „*mismatched framing*“, indem sie den Nutzen als relative Risikoreduktion und den Schaden als absoluten Risikoanstieg quantifizierten. Es ist davon auszugehen,

dass dies nicht intentional geschieht, da gut dokumentiert ist (s. oben), dass ÄrztInnen selbst nur selten zwischen relativen und absoluten Risiken unterscheiden können. Um sich selbst und PatientInnen vor einer solchen numerischen Verwirrung zu schützen, sollten Prozentzahlen (und ebenso auch das in medizinischen Journalen oftmals verwendete relative Maß der Odds Ratio) stets hinterfragt und die dazugehörigen absoluten Angaben zu Basisraten (z. B. Anzahl der an Brustkrebs verstorbenen Frauen) mit den entsprechenden Referenzgruppen (=jeweilige untersuchte Gesamtpopulation) in der Kontroll- und in der Screeninggruppe gesucht werden. Die Darstellung der Ergebnisse in Form einer Faktenbox (Abb. 1) hat sich dabei in der Patientenkommunikation als besonders unterstützend erwiesen [19].

Es sei an dieser Stelle angemerkt, dass in Deutschland zum Teil hoch emotionale Kontroversen über die Sinnhaftigkeit und die tatsächliche Größe des Nutzens des populationsbasierten Mammographie-Screenings geführt werden. Dabei ist ein Argument der Befürworter, das die Daten des Cochrane-Reviews mit rund 600.000 Frauen noch aus Zeiten stammen, in denen die Mammographie nicht so sensitiv und Tumorthérapien nicht so effektiv wie heute waren. Für das niederländische Mammographie-Screeningprogramm wurde 2017 eine aktuelle Auswertung der niederländischen Daten für den Zeitraum bis zum Jahr 2012 im *British Medical Journal* publiziert, mit Ergebnissen, die weit ernüchternder sind als die des Cochrane-Reviews [20]. Lediglich die über die Jahre hinweg verbesserten Therapien zeigten einen relevanten Einfluss auf die Brustkrebssterblichkeit. Für die Mammographie selbst war weder ein relevanter Einfluss auf die Brustkrebssterblichkeit noch auf die fortgeschrittenen Krebsstadien nachweisbar. Hingegen hat sich das Ausmaß an Überdiagnosen mit der Einführung von sensitiveren Screeningverfahren wie der digitalen Mammographie verschärft, so die Autoren der Studie. Warum sensitivere Testverfahren im Rahmen von populationsbasierten Screenings nicht zwangsläufig besser sind als weniger sensitive,

Gynäkologie 2018 · 51:370–379 <https://doi.org/10.1007/s00129-018-4199-3>
 © Der/die Autor(en) 2018. Dieser Artikel ist eine Open-Access-Publikation.

O. Wegwarth

Brustkrebsfrüherkennung – Nutzen und Risiken richtig kommunizieren

Zusammenfassung

Eine effiziente Gesundheitsversorgung braucht gut informierte Ärzte und Patienten. In unserem gegenwärtigen Gesundheitssystem gibt es beide nicht. Um das Problem zu illustrieren: In einer repräsentativen, europäischen Studie mit mehr als 5000 Frauen überschätzten 92 % den Nutzen der Mammografie um das 10- bis 200-Fache oder kannte ihn schlicht nicht. In einer Studie mit über 300 Menschen, die regelmäßig an einer oder mehreren Früherkennungsmaßnahmen teilnehmen, hatten über 90 % noch nie etwas vom größten Schaden der Früherkennung – Überdiagnose und Überbehandlung – von ihren ÄrztInnen gehört. Von 160 GynäkologInnen waren etwa 80 % nicht in

der Lage, den positiven Vorhersagewert der Mammographie zu errechnen. Warum ist die Risikokompetenz im Gesundheitssystem so gering? Die Forschung zeigt, dass es weniger kognitive Defizite sind, die Menschen davon abhalten, Evidenz zu verstehen, sondern vielmehr die Art und Weise, wie statistische Informationen präsentiert werden, so etwa durch verzerrte Berichterstattung in medizinischen Fachzeitschriften, Patienteninformationen und den Medien, die relative statt absolute Risiken und andere irreführende Statistik verwenden. Was kann getan werden? Medizinische Fakultäten sollten damit beginnen, ihren Studierenden simple Techniken der Risikokommunikation beizubringen,

die sie dabei unterstützen, medizinische Statistik richtig zu verstehen. Erwiesen ist, dass diese Techniken einen Großteil der statistischen Verwirrung eliminieren und zum informierten Entscheiden beitragen. Eine kritische Masse informierter Menschen wird nicht alle Probleme des Gesundheitssystems lösen, aber sie kann der auslösende Faktor für eine bessere Versorgung sein.

Schlüsselwörter

Krebsfrüherkennung · Risiko · Defensive Medizin · Unsicherheit · Gesundheitsversorgungsforschung

Early recognition of breast cancer—Communicate benefits and risks correctly

Abstract

An efficient health care requires both informed doctors and patients. Our current healthcare system falls short on both counts. Most doctors and patients do not understand the available medical evidence. To illustrate the extent of the problem in the setting of cancer screening: in a representative sample of some 5000 women in 9 European countries, 92% overestimated the reduction of breast cancer mortality by mammography by a factor of 10–200 or did not know. Of more than 300 US citizens who regularly attended one or more cancer screening tests, more than 90% had never been informed about the biggest harms of screening, overdiagnosis and overtreatment, by their physicians.

Among 160 German gynecologists, some 80% did not understand the positive predictive value of a positive mammogram. Why is risk literacy so scarce in health care? Research has demonstrated that the problem lies less in stable cognitive deficits than in how information is presented to physicians and patients. This includes biased reporting in medical journals, brochures, and the media that use relative risks and other misleading statistics. What can be done? Every medical school should teach its students how to understand evidence in general and health statistics in particular. To cultivate informed patients, elementary and high schools should start teaching the mathematics of

uncertainty and statistical thinking. Guidelines about complete and transparent reporting in journals, brochures, and the media need to be better enforced, and laws need to be changed in order to protect patients and doctors alike against the practice of defensive medicine instead of encouraging it. A critical mass of informed citizens will not resolve all healthcare problems, but it can constitute a major triggering factor for better care.

Keywords

Early detection of cancer · Risk · Defensive medicine · Uncertainty · Health services research

wird auch im nächsten Abschnitt besprochen.

Kommunizieren von Testergebnissen

Im Zuge des zunehmenden Angebots von Testungen in der Medizin kommt dem ärztlichen Verständnis von positiven und negativen Testergebnissen eine immer größere Bedeutung zu. So fragen sich die meisten Frauen nach Erhalt eines positiven Mammogramms, ob sie nun mit Sicherheit Brustkrebs haben bzw. wie wahrscheinlich es ist, dass ein Brust-

krebs vorliegt. Die Wahrscheinlichkeit, dass die Erkrankung bei einem positiven Testergebnis tatsächlich vorliegt, wird *positiver Vorhersagewert* („*positive predictive value*“, PPV) genannt. Sind ÄrztInnen in der Lage, Frauen auf diese wichtige Frage eine richtige Antwort zu geben? Zu Beginn einer Fortbildung stellten wir 160 GynäkologInnen mit allen erforderlichen Daten aus, derer es bedarf, um einer Frau die Antwort auf die Frage nach dem PPV zu geben. Die Daten selbst waren so aufbereitet, wie medizinische Studien normalerweise über die Eigenschaften eines diagnostischen Tests

berichten [5], und die Aufgabenstellung war wie folgt:

Nehmen Sie an, dass Sie in einer bestimmten Region ein Mammographie-Screening durchführen. Sie haben folgende Informationen über die Frauen in dieser Region:

- Die Wahrscheinlichkeit, dass eine Frau Brustkrebs hat, beträgt 1 % (Prävalenz).
- Wenn eine Frau Brustkrebs hat, liegt die Wahrscheinlichkeit eines positiven Tests bei 90 % (Sensitivität).
- Bei einer Frau, die keinen Brustkrebs hat, liegt die Wahrscheinlichkeit, dass

Brustkrebs-Früherkennung durch Mammographie-Screening

○ ○ ○ ○ HARDING-ZENTRUM FÜR
○ ○ ○ ○ RISIKOKOMPETENZ

Zahlen für Frauen ab 50 Jahren, die etwa 11 Jahre am Mammographie-Screening teilgenommen oder nicht teilgenommen haben.

	2.000 Frauen ohne Screening*	2.000 Frauen mit Screening*
Nutzen		
Wie viele Frauen starben an Brustkrebs?	11	10
Wie viele Frauen starben insgesamt an Krebs?	42	kein Unterschied
Schaden		
Wie viele Frauen erhielten fälschlicherweise ein positives Ergebnis und hatten zusätzliche Untersuchungen oder eine Gewebeentnahme (Biopsie)?	-	etwa 200
Wie viele Frauen mit nicht fortschreitendem Brustkrebs hatten eine unnötige teilweise oder vollständige Entfernung der Brust?	-	10

*Zur Erleichterung der Lesbarkeit wurden die Ergebnisse nach der Körnmastelle gerundet.

Kurz zusammengefasst: Mittels Mammographie-Screening konnte die Anzahl an Frauen, die an Brustkrebs starben, gesenkt werden. Allerdings hatte dies keinen Einfluss auf die Gesamtzahl an Frauen, die an Krebs starben. Von allen Frauen, die an dem Screening teilnahmen, wurden einige mit nicht fortschreitendem Krebs überdiagnostiziert und unnötig behandelt.

Quelle: Gøtzsche & Jørgensen. *Cochrane Database Syst Rev* 2013(6):CD001877.

Abb. 1 ▲ Faktenbox zur Darstellung des Nutzens und des Schadens der Mammographie. Im konkreten Fall wurden alle Endpunkte auf den Nenner von 1000 Frauen berechnet und die brustkrebspezifische Mortalität für Frauen im Sinne eines erleichterten Verständnisses aufseiten der PatientInnen gerundet. (Aus [10])

sie trotzdem positiv getestet wird, bei 9 % (Falsch-positiv-Rate = 100 – Spezifität).

Eine Frau wird positiv getestet. Sie will nun von Ihnen wissen, ob dies für sie mit Sicherheit „Brustkrebs“ bedeutet oder welche Wahrscheinlichkeit hierfür vorliegt. Wie lautet die beste Antwort?

- Die Wahrscheinlichkeit von Brustkrebs liegt bei etwa 81 %.
- Die Wahrscheinlichkeit von Brustkrebs liegt bei etwa 90 %.
- Die Wahrscheinlichkeit von Brustkrebs liegt bei etwa 10 %.
- Die Wahrscheinlichkeit von Brustkrebs liegt bei etwa 1 %.

Die GynäkologInnen konnten die Antwort entweder aus den ausgeteilten Gesundheitsstatistiken herleiten oder sich einfach an das erinnern, was sie eigentlich von vornherein hätten wissen müssen. In beiden Fällen ist die beste Antwortmöglichkeit „C“, also: Nur etwa eine von jeweils 10 Frauen mit positivem Früherkennungsergebnis hat tatsächlich Brustkrebs, die übrigen 9 Frauen erhalten ein falsch-positives Testergebnis. Nur 21 % der GynäkologInnen gaben die beste Antwort. Die Mehrheit (60 %) wählte fälschlicherweise die Antworten „90 %“ oder

„81 %“ und überschätzte damit die Wahrscheinlichkeit, dass eine positiv getestete Frau an Brustkrebs erkrankt ist, erheblich. Beunruhigend war weiterhin die große Bandbreite der Schätzungen, die von einer 1%igen bis zu einer 90%igen Wahrscheinlichkeit reichten.

Warum es ÄrztInnen so schwerfällt, anhand dieser Zahlen den *positiven Vorhersagewert* zu errechnen, verdeutlicht **Abb. 2a**. Um zu bestimmen, mit welcher Wahrscheinlichkeit eine Patientin bei Vorliegen eines positiven Mammographiebefundes tatsächlich Brustkrebs hat, wird die Formel des Bayes-Theorems (**Abb. 2a**) benötigt [21]. Nur wenigen Menschen ist diese Formel so eingängig, dass sie sich diese über den Vermittlungsmoment hinaus merken würden. Studien zeigen, dass es der Mehrzahl von ÄrztInnen trotz Verfügbarkeit der benötigten Kennzahlen nicht möglich ist, den PPV zu errechnen [22–24]. Ein einfacher Ausweg aus dem schwer eingängigen Bayes-Theorem ist die Verwendung von *natürlichen Häufigkeiten* [25]. Hierzu werden die entsprechenden prozentualen Informationen in Häufigkeiten überführt:

- **Prävalenz 1 %:** Von je 10.000 Frauen werden etwa 100 an Brustkrebs erkrankt sein.

- **Sensitivität 90 %:** Von diesen 100 tatsächlich erkrankten Frauen erhalten 90 ein positives Testergebnis.
- **Falsch-positiv-Rate 9 %:** Von den 9900 Frauen, die nicht an Brustkrebs erkrankt sind, erhalten 891 ebenfalls ein positives Testergebnis.

Diese Darstellung macht deutlich, dass bei einer Testung von 10.000 Frauen mit einer angenommenen Prävalenz von 1 % ca. 90 richtig-positive Testergebnisse und 891 falsch-positive Testergebnisse zu erwarten sind. Um den PPV nun final zu bestimmen, wird die Anzahl der richtig-positiven Testergebnisse (=90) ins Verhältnis gesetzt zur Gesamtheit aller positiven Testergebnisse (891 + 90 = 981), was in einem PPV von rund 10 % resultiert. Wie natürliche Häufigkeiten auch in grafischer Form, zum Beispiel zur unterstützenden Kommunikation von Testergebnissen mit PatientInnen, dargestellt werden können, zeigt **Abb. 2b**. Die Verwendung *natürlicher Häufigkeiten* zeigt erstaunliche Effekte beim Verständnis von Testergebnissen. Nicht nur ÄrztInnen und Medizinstudierende können damit auf einmal problemlos den positiven (oder auch den negativen) Vorhersagewert von Tests bestimmen, sondern auch Laien, Richter und 10-jährige Schulkinder [25–28].

Und zu guter Letzt lässt sich mithilfe der grafischen Darstellung noch ein weiterer wichtiger Umstand verdeutlichen: Besonders bei Krankheiten mit niedriger Prävalenz – wozu Brustkrebs in der Allgemeinbevölkerung gehört – ändert eine Verbesserung der Sensitivität des Testverfahrens (zielt immer auf tatsächlich Erkrankte) nichts am Vorkommen falsch-positiver Ergebnisse und damit kaum etwas an einem geringen positiven Vorhersagewert der Testung. So würde etwa eine Erhöhung der Sensitivität von 90 auf 99 % zwar dazu beitragen, dass von 100 tatsächlich Erkrankten pro 10.000 nun auch 99 korrekt entdeckt werden. Doch im Verhältnis zu den weiterhin bestehenden 891 falsch-positiv Diagnostizierten auf der Seite der Nichterkrankten verbleibt der PPV bei 10 %. Bei Erkrankungen mit niedriger Prävalenz ist deshalb eine hohe Spezifität entscheidender, auch besonders mit Blick auf die lang-

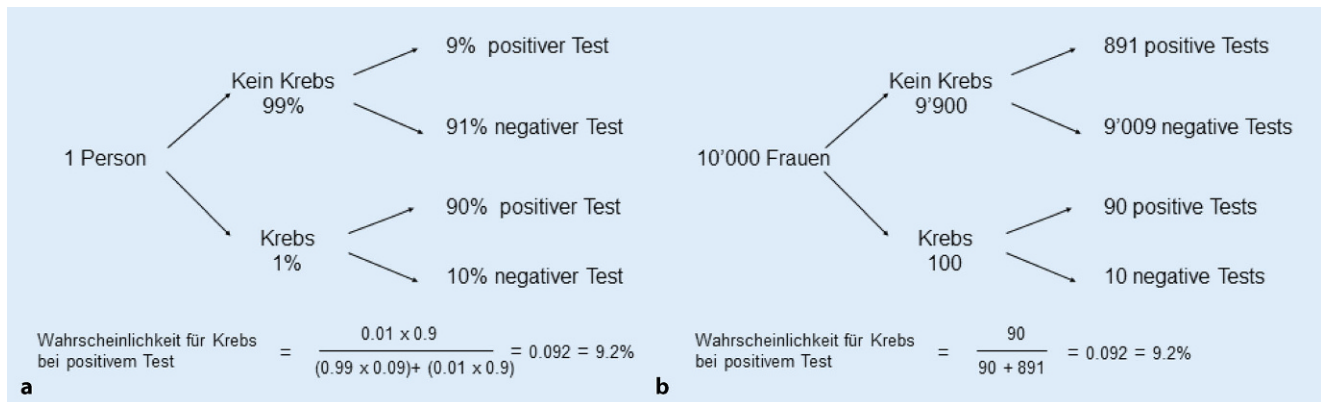


Abb. 2 ▲ Verschiedene Möglichkeiten der Berechnung der Wahrscheinlichkeit von Brustkrebs bei Vorliegen eines positiven Mammographiebefundes. **a** illustriert die Kalkulation mit bedingten Wahrscheinlichkeiten, ausgedrückt in Prozent, **b** zeigt, dass natürliche Häufigkeiten die Kalkulation vereinfachen. Relative Häufigkeiten, die numerisch mit bedingten Wahrscheinlichkeiten identisch sind, vereinfachen die Kalkulation dagegen nicht

fristigen psychischen Belastungserscheinungen, die nach falsch-positiven Befunden dokumentiert wurden: Noch Jahre später nach dem Erhalt eines falsch-positiven Befundes berichten Frauen von erhöhten brustkrebsbezogenen Ängsten, tasten vermehrt ihre Brust ab auf der Suche nach einem auffälligen Knoten und gehen aus Sorge vor Brustkrebs häufiger zu Früherkennungsuntersuchungen [29, 30]. Je näher die Spezifität an 100% gelangt, desto kleiner ist der Anteil an falsch-positiven Ergebnissen. Hätten wir in unserem Beispiel in **Abb. 2b** eine Spezifität von 99% und damit eine Falsch-positiv-Rate von 1%, wären bei den 9900 Frauen ohne Brustkrebserkrankung nur noch 99 statt 891 falsch-positive Ergebnisse zu erwarten. Würde man diese dann ins Verhältnis zu den 90 richtig-positiv Getesteten (**Abb. 2b**) setzen, dann erhielte man einen PPV von zumindest 48%.

» Das Problem der Überdiagnosen hat sich durch sensitivere Screeningverfahren noch verschärft

Liegt der Fokus eines Testverfahrens, das maßgeblich bei Niedrigprävalenzerkrankungen eingesetzt wird, eher auf einer hohen Sensitivität (zielt auf Erkrankte, davon gibt es bei Erkrankungen mit niedriger Prävalenz nur wenige) und nicht auf einer hohen Spezifität (zielt auf Nichter-

krankte, davon gibt es bei Erkrankungen mit niedriger Prävalenz viele), kann daraus noch ein weiteres Problem resultieren: mehr Überdiagnosen. Je sensitiver ein Testverfahren ist, umso mehr „reagiert“ dieses auch auf jene Zellveränderungen, die zwar letztlich der pathologischen Definition eines Tumors entsprechen, aber perspektivisch weder Symptome noch krebspezifischen Tod verursachen. Wie im vorigen Abschnitt bereits erwähnt, zeigen jüngste Studien, dass sich das Problem der Zunahme von Überdiagnosen mit der Einführung sensitiverer Screeningverfahren wie der digitalen Mammographie verschärft hat.

» Zwischen Sensitivität und Spezifität ist in der Testentwicklung immer nur ein Trade-off zu erreichen

Einer hohen Sensitivität kommt aber sehr wohl eine hohe Relevanz zu, wenn das Testverfahren in einer Hochrisikogruppe eingesetzt wird. Um das zu verdeutlichen: Nehmen wir an, dass die Mammographie in einer Population von Frauen eingesetzt wird, in der das Brustkrebsrisiko bei 80% liegt. In diesem Falle kämen auf 10.000 getestete Frauen 8000 mit Brustkrebs, wovon bei einer 90%igen Sensitivität 7200 auch korrekt positiv testen würden. Von den 2000 Frauen, die nicht an Brustkrebs erkrankt sind, erhielten bei einer Falsch-positiv-Rate von 9% noch

rund 180 Frauen ein falsch-positives Ergebnis. Von den dann insgesamt 7380 positiven Ergebnissen wären 7200 richtig positive und man erhielte einen PPV von rund 98%.

Ohne Frage wäre es wünschenswert, für beide Testgütekriterien (Sensitivität und Spezifität) möglichst hohe Werte zu erlangen. Doch zwischen diesen beiden Kriterien ist in der Testentwicklung immer nur ein Trade-off zu erreichen, das heißt ein sehr hoch-sensitiver Test kann nicht gleichzeitig auch eine sehr hohe Spezifität haben und umgekehrt. Um dem Problem der Überdiagnose und der falsch-positiven Befunde habhaft und dem gleichzeitigem Wunsch nach rechtzeitiger, richtiger Diagnosestellung gerecht zu werden, richten Forschungsaktivitäten im Bereich der Früherkennung weiblicher Tumoren zunehmend den Blick auf das Potenzial von risikostratifizierter Früherkennung. Doch bis erste Daten zum Nutzen einer solchen Risikostratifizierung vorliegen werden, sollte man sich bei der Beurteilung der Aussagekraft von Testergebnissen weiterhin vergegenwärtigen, ob die Testung hochsensitiv oder eher hochspezifisch ist und ob in einer Hochrisikogruppe oder in der Allgemeinbevölkerung getestet wird.

Überlebensraten im Kontext von Früherkennung

Jedes Jahr im Oktober ist in Amerika und mittlerweile auch in Deutschland der

Tab. 1 Angaben zur Aufgabe für MedizinerInnen, sich für oder gegen den Rat zur Teilnahme ihrer PatientInnen an einem Screeningprogramm zu entscheiden. (Aus [34], klinische Studie zum Verständnis medizinischer Statistiken)

Szenario 1	Szenario 2
Fünfjahresüberlebensrate für Tumor X	Mortalitätsrate für Tumor Z über 5 Jahre
Ohne Screening: 68%	Ohne Screening: 2 pro 1000
Mit Screening: 99%	Mit Screening: 1,6 pro 1000

„Breast cancer awareness month“. Keine Organisation ist in Amerika mehr prominent in der Bewerbung dieses Monats als Susan G. Komen®, die weltgrößte Brustkrebsstiftung und Erfinderin des „pink ribbon“. Ihr Portfolio an Aktivitäten beinhaltet eine Vielzahl ehrenwerter Aktivitäten, die dazu beitragen sollen, Frauen vor dem Brustkrebstod zu bewahren, Patientinnen in ihren Rechten zu stärken, Lebens- wie Versorgungsqualität sicherzustellen und Forschung zu noch besseren Krebstherapien zu motivieren. Die Stiftung ist jedoch auch dafür bekannt, das Mammographiescreening zu bewerben [31]. Auf einem ihrer Werbeplakate, die 2011 während des „Breast cancer awareness month“ in ganz Amerika aufgestellt wurden, stand [31]:

What is the key to surviving breast cancer? YOU. Early detection saves lives. The 5-year survival of breast cancer when caught early is 98%. When it's not? 23%.¹

Der hier propagierte Unterschied im Überleben zwischen Frauen, die an der Mammographie teilnehmen, und denen, die es nicht tun, wirkt mit 75% eindrucksvoll. Doch gibt es diesen Unterschied überhaupt? Nein. Das werden viele Frauen aber wahrscheinlich nie erfahren. Unterschiede in Überlebensraten haben im Kontext von Früherkennungen keinerlei Korrespondenz zu Unterschieden in der Sterblichkeit zwischen Menschen, die an Früherkennung teilnehmen bzw. nicht teilnehmen [32]. Warum ist das so? Der Grund dafür ist die Art der Berechnung der Überlebensrate im Kontext von Früherkennung. Dabei steht im Nenner der Statistik die

Anzahl aller mit Krebs Diagnostizierten und im Zähler die Anzahl der Diagnostizierten, die nach einem bestimmten Zeitraum, etwa nach 5 oder 10 Jahren, noch leben. Das Problem besteht nun darin, dass der Nenner aufgrund der maßgeblichen Eigenschaft der Früherkennung – nämlich einen potenziellen Tumor schon deutlich vor dem Verursachen von Symptomen zu entdecken – künstlich aufgebläht wird, was zu 2 systematischen Bias führt. Der erste heißt Vorlaufzeit-Bias („lead time bias“). Um diesen zu verstehen, stellen Sie sich 100 Menschen vor, die nicht an der Früherkennung teilnehmen. Stellen Sie sich weiter vor, dass bei allen mit 67 Jahren ein Tumor aufgrund von Symptomen entdeckt wird und alle mit 70 Jahren daran verstorben sind. Die Fünfjahresüberlebensrate wäre in diesen Falle 0%. Nun stellen Sie sich vor, dass diese 100 Menschen alle an der Früherkennung teilnehmen. Durch die Eigenschaft der Früherkennung, den Tumor schon weit vor dem Entstehen von Symptomen zu detektieren, erhalten nun alle 100 Personen im Alter von bereits 60 Jahren ihre Diagnose, sind aber wieder alle mit 70 Jahren verstorben. In diesem Falle wäre die Fünfjahresüberlebensrate 100%. Obwohl die Teilnahme an der Früherkennung zu keinem zusätzlichen Jahr an Leben führte, sondern lediglich zu einer verlängerten Zeit der Diagnose, ist der Unterschied im Fünfjahresüberleben zwischen den beiden Gruppen dennoch 100%. Der zweite Bias, der zur künstlichen Inflation der Überlebensrate zum Vorteil der Früherkennungsgruppe führt, ist der Überdiagnose-Bias („overdiagnosis bias“). Um diesen Bias zu verstehen, stellen Sie sich 10.000 Menschen vor, die nicht an der Früherkennung teilnehmen. Nehmen Sie weiter an, dass innerhalb dieser Population 100 Menschen einen tatsächlich progressiven

Tumor haben, der bei allen 100 Menschen anhand von Symptomen entdeckt wird. Fünf Jahre nach Diagnosestellung durch Symptome leben noch 68 Menschen und 32 Personen sind an dem Tumor verstorben. Die entsprechende Fünfjahresüberlebensrate liegt damit bei 68%. Stellen Sie sich nun vor, dass die besagte Population von 10.000 Menschen zur Früherkennung geht. Da durch die Früherkennung auch nichtprogressive Tumoren entdeckt werden, die per definitionem keinerlei Symptome oder krebspezifischen Tod gebracht hätten, werden auch diese neben den tatsächlich progressiven Tumoren in den Nenner der Überlebensstatistik eingeschlossen. Nehmen wir an, dass neben den 100 tatsächlich progressiven Tumoren nun auch 100 nichtprogressive Tumoren durch die Früherkennung entdeckt würden, die definitiv nicht zu dem tumorspezifischen Tod in den nächsten 5 Jahre führen würden. Diese 100 werden nun zu den 68 Überlebenden mit tatsächlich progressivem Tumor im Zähler dazugerechnet (= 168) und ebenfalls im Nenner zu den 100 Diagnostizierten mit progressivem Tumor dazu addiert (100 + 100 = 200). Dadurch ergibt sich nun eine Fünfjahresüberlebensrate von 84%, ohne dass aber tatsächlich ein Mensch mehr gerettet wurde, da die 100 Menschen mit nichtprogressiven Tumoren so oder so überlebt hätten. Die Überdiagnose in der Früherkennungsgruppe ist auch der Grund dafür, dass mehr Tumoren in der Früherkennungsgruppe entdeckt werden (definitiv kein Zeichen für Erfolg, sondern im Gegenteil: für Schaden) und dass das relative Verhältnis von „prognostisch günstigen“ und „prognostisch ungünstigen“ Tumoren in der Früherkennungsgruppe scheinbar besser ist. Ob das Entdecken von mehr und früheren Stadien tatsächlich klinische Relevanz hat bzw. ein Erfolg der Früherkennung ist, kann tatsächlich nur durch eine Statistik unverzerrt beantwortet werden: durch eine Reduktion der Mortalitätsrate. Im Gegensatz zur Überlebensrate hat die Mortalitätsrate die jeweilige Gesamtpopulation im Nenner (z. B. alle 254.987 Frauen der Früherkennungsgruppe, alle 254.678 Frauen in der Kontrollgruppe) und die absolute tumorspezifische Mor-

¹ Auf Deutsch etwa: „Was ist der Schlüssel dazu, Brustkrebs zu überleben? DU. Früherkennung rettet Leben. Die Fünfjahresüberlebensrate von Brustkrebs, wenn er früh erkannt wird, ist 98%. Wenn nicht? 23%.“

talität der jeweiligen Gruppe im Nenner und ist damit nicht durch die Art der Diagnosestellung beeinflusst.

Klinisch relevant wäre für die Beurteilung der Früherkennung auch noch eine Reduktion der Invasivität der Therapie, die einer Diagnose folgt. Bisherige randomisierte, kontrollierte Studien zu verschiedenen Früherkennungsmaßnahmen bleiben den Beweis für weniger invasive Therapie in der Früherkennungsgruppe im Vergleich zur Kontrollgruppe jedoch schuldig.

Wenn Menschen nun schon durch Patienteninformationen in die Irre geführt werden, können sie zumindest darauf hoffen, dass ÄrztInnen wissen, dass Überlebensraten ungeeignet zur Bewertung von Screeningmaßnahmen sind? Um diese Frage zu beantworten, wurden 65 FachärztInnen für Innere Medizin und für Urologie in verschiedenen Szenarien mit Fünfjahresüberlebensraten bzw. mit krankheitsspezifischen Mortalitätsraten konfrontiert. Den ÄrztInnen wurde dabei verschwiegen, dass beide den Effekt derselben Früherkennungsmaßnahme abbildeten, den der Prostatakrebsfrüherkennung [33]. Sechundsiebzig Prozent der MedizinerInnen waren bereit, einem fiktiven Patienten das Screening zu empfehlen, nachdem sie die entsprechenden Fünfjahresüberlebensraten gesehen hatten, 79% beurteilten die Früherkennung als effektiv. Nach der Präsentation der Mortalitätsraten empfahlen jedoch nur noch 8% derselben ÄrztInnen das Screening, und lediglich 5% beurteilten es als effektiv. Ferner sollten die befragten ÄrztInnen schätzen, wie viele Todesfälle pro 1000 Personen sich durch die regelmäßige Teilnahme am Screening vermeiden ließen. Die durchschnittliche Schätzung lag nach der Konfrontation mit den Fünfjahresüberlebensraten bei 150 verhinderten Todesfällen pro 1000 Personen. Die tatsächliche absolute Reduktion der krankheitsspezifischen Mortalität lag allerdings bei weniger als 1/1000. Erst nachdem die Mortalitätsraten gezeigt wurden, gelang es der Mehrzahl der ÄrztInnen, den tatsächlichen Nutzen des Screenings aus der Statistik abzuleiten. Nur 2 waren in der Lage, den Begriff „Vorlaufzeit-Bias“ zu erklären, keiner konnte etwas mit dem

Begriff „Überdiagnose-Bias“ anfangen. Eine ähnliche Studie mit 412 amerikanischen ÄrztInnen zeigte vergleichbare Ergebnisse. In dieser Studie lautete die Aufgabenstellung:

Stellen Sie sich vor, dass Ihr 55-jähriger Patient Sie um Rat bezüglich der Früherkennung für den Tumor X/Z bittet. Ihnen liegen die in Tab. 1 aufgeführten Daten für Patienten zwischen 50 und 69 Jahren vor, die aus einer großen randomisierten Studie mit 10-jähriger Laufzeit stammen. Würden Sie Ihrem Patienten anhand dieser Daten die Teilnahme am Screening empfehlen?

Wurde der Nutzen in dieser amerikanischen Studie durch Fünfjahresüberlebensraten propagiert, empfahlen 69% aller ÄrztInnen ihren PatientInnen die Früherkennung, jedoch nur 23%, wenn der Nutzen anhand von Mortalitätsraten beschrieben war. Danach befragt, welche Statistik (Überlebensraten vs. Mortalitätsraten) beweisen würde, dass Screening Leben rette, waren ÄrztInnen nicht in der Lage, zwischen der invaliden (Überlebensrate) und der validen (Mortalitätsrate) zu unterscheiden: 79% nahmen (fälschlicherweise) an, dass Überlebensraten das beweisen können, und zu derselben Zeit nahmen 81% desselben Kollektivs dies (richtigerweise) für Mortalitätsraten an. Ferner glaubten 47% inkorrekt, mehr entdeckte Tumoren in der Früherkennungsgruppe würden beweisen, dass mehr Leben in der Früherkennungsgruppe gerettet werden würden.

» Die Überlebensrate ist eine invalide Statistik im Kontext von Früherkennung

Die Überlebensrate ist eine invalide Statistik im Kontext von Früherkennung, und mehr entdeckte Tumoren bzw. mehr früh entdeckte Tumoren in der Früherkennungsgruppe sind meist ein Hinweis für Überdiagnosen (Schaden) und nicht der Beweis für ein Mehr an Nutzen. Die Verwendung dieser Statistiken in der Kommunikation weckt unrealistische Erwartungen nicht nur bei PatientInnen, sondern auch bei ÄrztInnen.

Wenn das Ziel die informierte Entscheidung und nicht das Überreden von PatientInnen ist, müssen PatientInnen stattdessen Zahlen darüber erhalten, wie viele PatientInnen pro einer gewissen Gruppengröße mit und ohne Screening an dem entsprechenden Tumor versterben (absolute Mortalitätsrate). Es sei hier betont, dass Überlebensraten lediglich im Kontext von Früherkennung invalide Statistiken sind. Im Kontext der Bewertung von Behandlungen sind sie eine valide Statistik, da sie in diesem Zusammenhang nicht durch die Art der Diagnosestellung beeinflusst sind.

Schlussfolgerungen

Jede medizinische Information kann auf 2 Arten dargestellt werden: eine, die für die Mehrzahl der ÄrztInnen und PatientInnen irreführend ist, und eine, die transparent und leicht verständlich ist. Entgegen der Annahme, dass es nur PatientInnen wären, die Schwierigkeiten beim Verstehen von medizinischer Statistik haben, lassen sich auch viele ÄrztInnen von Fachstatistiken in die Irre führen. Sie sind damit außerstande, die Nutzen-Schaden-Bilanz angebotener medizinischer Maßnahmen realistisch einzuschätzen und PatientInnen die nötigen Fakten zu benennen, derer es für eine informierte Entscheidung zur Früherkennung bedürfte. Intransparente Statistiken in medizinischen Fachzeitschriften, irreführende Informationen in Broschüren für ÄrztInnen und PatientInnen sowie mangelnde Schulung in medizinischer Statistik und Risikokommunikation an medizinischen Fakultäten sind wichtige Gründe für die fehlende Risikokompetenz vieler ÄrztInnen.

Was kann gegen diese Situation getan werden? Medizinische Fakultäten sollten damit beginnen, ihren Studierenden simple Techniken der Risikokommunikation beizubringen, die sie dabei unterstützen, Evidenz im Allgemeinen und medizinische Statistik im Speziellen richtig zu verstehen. Solche – in diesem Artikel vorgestellte – Techniken der Risikokommunikation haben in Studien gezeigt, dass sie einen Großteil der statistischen Verwirrung aufseiten der ÄrztInnen (und PatientInnen) eliminie-

ren und damit zu einer informierten Entscheidung beitragen. Eine kritische Masse informierter Menschen wird nicht die Gesamtheit der Probleme unseres Gesundheitssystems lösen, aber sie kann der auslösende Faktor für eine bessere Versorgung sein.

Fazit für die Praxis

- Relative Risikoangaben führen zu massiven Überschätzungen des Nutzens von medizinischen Maßnahmen.
- Nutzen *und* Schaden sollten immer als absolutes Risiko dargestellt werden.
- Das Verständnis wird unterstützt, wenn die absoluten Risiken (z. B. Mortalität durch Brustkrebs) für die Kontroll- und für die Früherkennungsgruppe auf denselben Nenner (z. B. pro 1000 Frauen) adjustiert werden, da sich die Gruppengrößen in Studien oft unterscheiden.
- Zur Berechnung des PPV sollten die benötigte Prävalenz der Erkrankung, die Sensitivität und die Spezifität des Tests in natürliche Häufigkeiten umgewandelt werden.
- Bei Erkrankungen mit niedriger Prävalenz (<10 %) kommt der Spezifität mit Blick auf das Risiko falsch-positiver Diagnosen eine besondere Rolle zu.
- Die Fünfjahresüberlebensrate ist im Kontext von Früherkennung eine invalide statistische Angabe.
- Auch das Entdecken von mehr Tumoren bzw. mehr Tumoren in früheren Stadien in der Früherkennungsgruppe beweist nicht den Nutzen von Screening. Der Beweis kann nur durch eine Reduktion der Mortalitätsrate erbracht werden.

Korrespondenzadresse



PD Dr. O. Wegwarth
Max-Planck-Institut für
Bildungsforschung, Center for
Adaptive Rationality/Harding
Center for Risk Literacy
Lentzeallee 94, 14195 Berlin,
Deutschland
wegwarth@mpib-
berlin.mpg.de

Funding. Open access funding provided by Max Planck Society.

Einhaltung ethischer Richtlinien

Interessenkonflikt. O. Wegwarth gibt an, dass kein Interessenkonflikt besteht.

Für den hier vorliegenden Beitrag wurden vom Autor keine Studien an Menschen oder Tieren durchgeführt. Für die im Beitrag aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

Open Access. Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Literatur

1. Sirovich BE, Schwartz LM, Woloshin S (2003) Screening men for prostate and colorectal cancer in the United States: does practice reflect the evidence? *JAMA* 289(11):1414–1430
2. Gigerenzer G, Gray JAM (2011) Launching the century of the patient. In: Gigerenzer G, Gray JAM (Hrsg) *Better doctors, better patients, better decisions: envisioning healthcare 2020*. Strüngmann forum report 6. MIT Press, Cambridge, S 1–19
3. Schwartz LM, Woloshin S, Fowler FJ, Welch HG (2004) Enthusiasm for cancer screening in the United States. *JAMA* 291:71–78
4. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S (2007) Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest* 8:53–96
5. Gigerenzer G, Gray JAM (2012) *Better doctors, better patients, better decisions: envisioning healthcare 2020*. Strüngmann forum report. MIT Press, Cambridge
6. Sunstein CR, Thaler RH (2003) Libertarian paternalism is not an oxymoron. *Univ Chic Law Rev* 70:1159–1202
7. Nyström L, Larsson LG, Wall S, Rutqvist LE, Andersson I, Bjurström N et al (1996) An overview of the Swedish randomised mammography trials: total mortality pattern and the representativity of the study cohorts. *J Med Screen* 3:85–87
8. Nyström L, Andersson I, Bjurström N, Frisell J, Nordenskjöld B, Rutqvist LE (2002) Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 359:909–919
9. Gøtzsche P, Nielsen M (2006) Screening for breast cancer with mammography. *Cochrane Database Syst Rev*. <https://doi.org/10.1002/14651858.cd001877.pub2>
10. Gøtzsche PC, Jørgensen KJ (2013) Screening for breast cancer with mammography. *Cochrane Database Syst Rev*. <https://doi.org/10.1002/14651858.cd001877.pub5>
11. Wegwarth O, Gigerenzer G (2011) “There is nothing to worry about”: gynecologists’ counseling on mammography. *Patient Educ Couns* 84(2):251–256
12. Gigerenzer G, Mata J, Frank R (2009) Public knowledge of benefits of breast and prostate cancer screening in Europe. *J Natl Cancer Inst* 101(17):1216–1220
13. Schüssler B (2005) Im Dialog: Ist Risiko überhaupt kommunizierbar, Herr Prof. Gigerenzer? *Frauenheilkd Aktuell* 14:25–31
14. Naylor CD, Chen E, Strauss B (1992) Measured enthusiasm: Does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 117:916–921
15. Ghosh AK, Ghosh K (2005) Translating evidence-based information into effective risk communication: current challenges and opportunities. *J Lab Clin Med* 145:171–180
16. Bramwell R, West H, Salmon P (2006) Health professionals’ and service users’ interpretation of screening test results: experimental study. *BMJ* 333:284–286
17. Sedrakyan A, Shih C (2007) Improving depiction of benefits and harms: analyses of studies of well-known therapeutics and review of high-impact medical journals. *Med Care* 45:523–528
18. Gigerenzer G, Wegwarth O, Feufel M (2010) Misleading communication of risk: editors should enforce transparent reporting in abstracts. *BMJ* 341:791–792
19. Schwartz LM, Woloshin S, Welch HG (2009) Using a drug facts box to communicate drug benefits and harms. *Ann Intern Med* 150(8):516–527
20. Autier P, Boniol M, Koechlin A, Pizot C, Boniol M (2017) Effectiveness of and overdiagnosis from mammography in the Netherlands: population based survey. *BMJ* 359:j5224
21. Gigerenzer G (2002) *Reckoning with risk: learning to live with uncertainty*. Penguin, London (US version: *Calculated risks*, Simon & Schuster)
22. Labarge AS, McCaffrey RJ, Brown TA (2003) Neuropsychologists’ ability to determine the predictive value of diagnostic tests. *Arch Clin Neuropsychol* 18:165–175
23. Steckelberg A, Balgenorth A, Berger J, Mühlhauser I (2004) Explaining computation of predictive values: 2 x 2 table versus frequency tree. A randomized controlled trial [ISRCTN74278823]. *BMC Med Educ*. <https://doi.org/10.1186/472-6920-4-13>
24. Eddy DM (1982) Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A (Hrsg) *Judgment under uncertainty: heuristics and biases*. Cambridge University Press, Cambridge, S 249–267
25. Hoffrage U, Gigerenzer G (1995) How to improve Bayesian reasoning without instruction: frequency formats. *Psychol Rev* 102:684–704
26. Zhu L, Gigerenzer G (2006) Children can solve Bayesian problems: the role of representation in mental computation. *Cognition* 98:287–308
27. Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G (2001) Response to Brian Butterworth: “Statistics: What seems natural?”. *Science* 292:853–855
28. Lindsey S, Hertwig R, Gigerenzer G (2003) Communicating statistical DNA evidence. *Jurimetrics* 43:147–163
29. Brewer NT, Salz T, Lillie SE (2007) The long-term effects of false-positive mammograms. *Ann Intern Med* 146:502–510
30. Salz T, Richman AR, Brewer NT (2010) Meta-analyses of the effect of false-positive mammograms

on generic and specific psychosocial outcomes. *Psychooncology* 19:1026–1034

31. Woloshin S, Schwartz LM (2012) How a charity oversells mammography. *BMJ* 345:e5132
32. Welch HG, Schwartz LM, Woloshin S (2000) Are increasing 5-year survival rates evidence of success against cancer? *JAMA* 283(22):2975–2978
33. Wegwarth O, Gaismaier W, Gigerenzer G (2011) Deceiving numbers: survival rates and their impact on doctors' risk communication. *Med Decis Making* 31(3):386–394
34. Wegwarth O, Schwartz LM, Woloshin S, Gaismaier W, Gigerenzer G (2012) Do physicians understand cancer screening statistics? A national survey of primary care physicians in the U.S. *Ann Intern Med* 156:340–349

N. Brochmann, E. Støkken Dahl

Viva la vagina! Alles über das weibliche Geschlecht

Das Buch hält mehr, als der Titel verspricht

Frankfurt (Main): **S. Fischer Verlag GmbH 2018, 1. Auflage, 400 S., (ISBN: 978-3103973389), broschiert 16,99 EUR**



Der Titel „Viva la vagina!“ macht neugierig – macht aber auch skeptisch. „Noch ein Buch aus dem Selbsthilferepertoire für Frauen mit Zweifeln und Unsicherheiten be-

züglich ihrer normalen körperlichen und sexuellen Ausstattung“?

Erfreulicherweise stellte sich dieses Buch zweier junger norwegischer Ärztinnen über den weiblichen Unterleib als besonders informativ, verständlich und unpräventiös, ja als lesenswertes Sachbuch heraus. Sehr schnell merkt man, dass es hier nicht etwa vornehmlich um die Stilisierung eines „lange vernachlässigten Organs“ geht, sondern dass die Lektüre dieses Buches die beste Idee ist, um den defizitären Kenntnissen von Frauen bezüglich ihres eigenen Körpers etwas Substantielles entgegenzusetzen. Und man versteht, warum dieses Buch bereits in 30 Länder verkauft wurde und jetzt im März 2018 auch in Deutschland erschienen ist. Mit angemessener Offenheit, ohne falsche Scham, mit gut dosierter Lockerheit und vor allem kenntnisreich und im Urteil ausgewogen wird das ganze Spektrum weiblicher Körperlichkeit und Sexualität thematisiert: Der Genitalbereich mit Vulva, Vagina, Klitoris, Hymen und den inneren Geschlechtsorganen ebenso wie Ausfluss, Menstruation, PMS (hier als Potentielles Mordsyndrom übersetzt), der ewige Kreislauf der Hormone und die Möglichkeiten der Konzeption werden kurzweilig und in lockerem Umgangston erklärt.

Besonders differenziert befassen sich die Autorinnen mit den sexuellen Fragen zum 1. Mal, zu den geschlechtstypischen Unterschieden männlichen und weiblichen sexuellen Begehrens, zu der Frage des vaginalen oder klitoridalen Orgasmus und warum es in unserem Leben und in unseren Beziehungen mit der sexuellen Lust auf und ab geht. Nie findet man den eventuellen Anfangsverdacht

bestätigt, dass hier vielleicht die Scham- und Tabugrenzen der Leserschaft (und vielleicht auch die eigenen) getestet werden sollen. Im Gegenteil: Leserinnen und vor allem die mit eher responsivem sexuellen Verlangen werden sich nach der Lektüre nicht länger als „wandelnde Abweichung“ vom Mainstream einer sexualisierten Gesellschaft verstehen müssen.

Neben der Sachkenntnis der Autorinnen kommt auch deren engagierter Schreibstil insbesondere dem Kapitel zur Empfängnisverhütung zugute. Die Fragen zur hormonellen und nichthormonellen Verhütung und zur Notfallverhütung sowie die damit verknüpften Bedenken vieler Frauen hinsichtlich eventueller Nebenwirkungen („Ist das nicht gefährlich?“ „Ist das sicher?“) werden in zur Nachahmung empfohlener Weise beantwortet.

Im letzten Kapitel stellen die Autorinnen unter dem Titel „Trouble im Intimbereich“ die gängigen gynäkologischen Befindlichkeitsstörungen und Krankheitsbilder inklusive der STI vor. Auf Augenhöhe mit der Leserin werden den Fachbegriffen PCO, Myome, Endometriose, Vulvodynie und („tropf, tropf, tropf“) auch der Inkontinenz Fakten entgegengehalten, die jeder Patientin einleuchten werden. Selbst das Thema Karzinom, dass man üblicherweise mit unzähligen (Schicksals-)Berichten, Fachliteratur oder Konferenzen verbindet, findet hier eine für Laien angemessene und spannende Bearbeitung. Nach dem Vorbild von „Darm mit Charme“ von Giulia Enders legen hier die Ärztinnen und Autorinnen Nina Brochmann und Ellen Støkken Dahl mit dem Buch „Viva la vagina!“ ein sehr differenziertes und im Stil junges und frisches „Aufklärungsbuch“ vor für alle Frauen (und deren Männer) - seine Existenz war überfällig. Ein großartiges Buch auch für alle Gynäkologinnen und Gynäkologen, die daran interessiert sind, wie ihre Patientinnen wirklich ticken und wie man mit ihnen kommunizieren sollte.

G. Gille (Lüneburg)