

Chances and risks in medical risk communication

Abstract

Communication between physicians and patients in everyday life is marked by a number of disruptive factors. Apart from specific interests, mistakes, and misunderstandings on both sides, there are main factors that contribute to the risk in risk communication. Using the example of mammography screening, the current work demonstrates how the meaning of test results and the informative value of measures taken to reduce risk are often misunderstood. Finally, the current work provides examples of successful risk communication.

Keywords: risk communication, mammography screening, Bayes' theorem, probabilities, natural frequencies, relative risk reduction, survival rates

Ulrich Hoffrage¹

Michael Koller²

1 Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland

2 Centre for Clinical Studies, University Hospital Regensburg, Regensburg, Germany

Patients and physicians sit in the same boat

Everyone knows the advice “*As regards to risks and side effects, ask your doctor or pharmacist*”. Strictly speaking, this sentence does not just constitute a good piece of advice for the readers of drug advertisements but, instead, is used by pharmaceutical companies to protect themselves against possible damage claims. As this demonstrates, risk and side effects are bad, everyone wants to avoid them, and risk communication can have various facets.

On the one hand, people do not want to become patients, and patients do not want to develop any side effects of drugs. On the other hand, physicians and pharmaceutical companies want to avoid making mistakes and being misunderstood or sued. Yet, risk avoidance is not the only common aim. Both sides are interdependent: While patients need the knowledge provided by experts, the different players within the health care system are not only there for the patients but also because of them. In other words, without patients, they would not have anything to do, and their professions would not even exist. Furthermore, each side wants to trust the other side: Distressed patients place their trust in physicians before the treatment, particularly because physicians usually provide the expected help, and physicians may rely on the gratefulness of their cured patients.

Although patients do not always get better after treatment, the commonalities continue even then. Sometimes, physicians make mistakes or do not exclusively act in the best interest of the patient because of constraints or conflicts of interest. Because many patients are aware of this problem, the trust they associate with physicians is sometimes mixed with a pinch of mistrust that is often intensified by the respective communication: regarding

risks and side effects when interacting with physicians, patients often ask other patients. But also patients make mistakes, for example, by ignoring instructions or advice and by unfairly holding the physician responsible for the deterioration of their condition, which leads to legal disputes from time to time. For this reason, also physicians have reason to report on risks and side effects when dealing with patients.

Risks of risk communication

The commonality between physicians and patients, that is, the possibility to help or hurt each other, can also be observed in risk communication. Laypeople are often unable to assess how dangerous their condition is and what could be done about it. Experts are able to help in this respect by providing the relevant information. This is the great chance of risk communication [1]. However, beware of the statement made by Benjamin Franklin “In this world nothing can be said to be certain, except death and taxes”. Even information provided by experts may be incorrect. Sometimes, such information is just incomplete or ambiguous, and it is not always clear if the physician or the patient is to be held responsible for the misunderstanding. Mistakes and possible misunderstandings, such as these, are the risks of risk communication.

Ideally, the knowledge divide between physicians and patients leads to the transparent transfer of relevant information. The situation in daily life, however, is often different because risk communication is not free of disruptive interests [2]. On the one hand, medical care can only be provided when costs are covered. On the other hand, it is legitimate and fully understandable if physicians and caregivers as well as authorities or hospital managements want to protect themselves against legal disputes. Additionally, the manner of risk communication

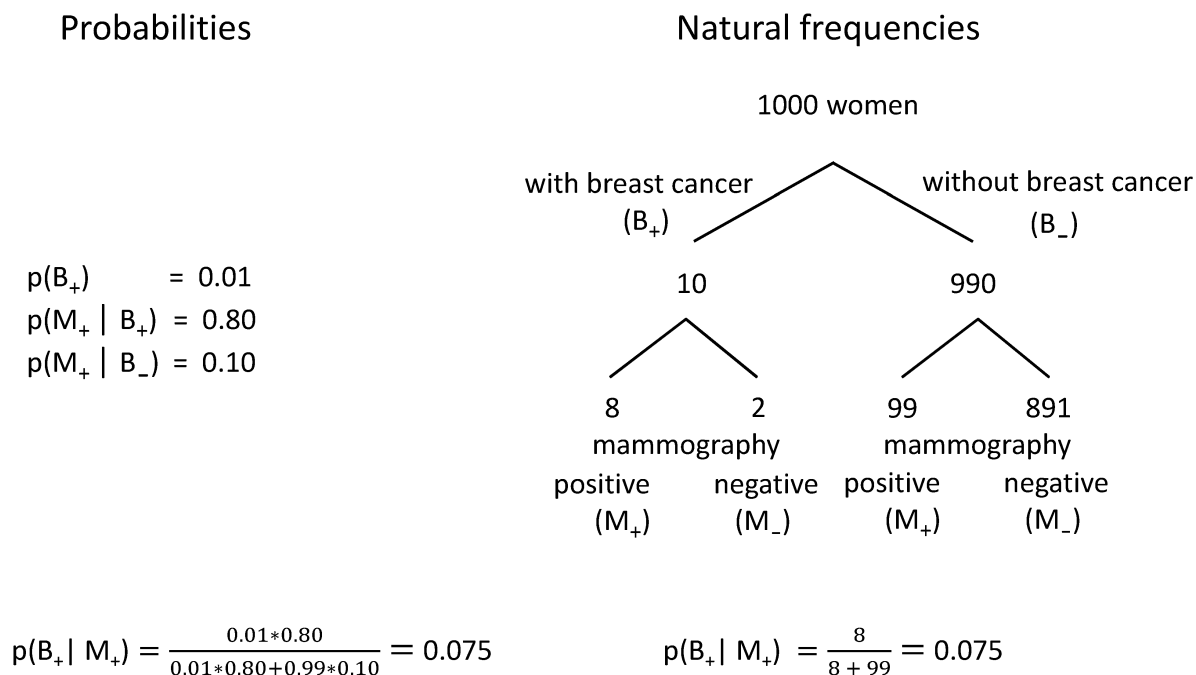


Figure 1: Representation of the same information in terms of probabilities and natural frequencies (adapted from [30]).

is also influenced by institutional aspects. In our opinion, the perhaps most serious challenge in transparent and understandable risk communication is the inability of some advisors to understand and convey figures correctly. We will present three different examples of how certain values and statistics that are commonly used in risk communication can be misunderstood and show how the risk of such misunderstandings could be reduced.

The meaning of test results

A 52-year-old woman accepted an invitation to undergo mammography screening. Despite the absence of relevant symptoms, the woman received a suspicious finding – requiring further examinations to clarify if the suspicious lump was indeed breast cancer. In 2011, about 4.9 million women aged between 50 and 69 years received an invitation to undergo mammography screening in Germany. Of the 2.7 million women accepting this invitation, 8.6% of the about 800,000 women who were screened for the first time and 3.2% of the about 1.9 million women for whom this was not their first mammography were asked to present for further examination [3].

However, what is the exact meaning of such a re-invitation? The most frequent reason is a positive finding at the first mammography screening. How high is the probability that it is really breast cancer? This probability, the so-called positive predictive value (PPV), depends on three different parameters. In a survey conducted in the United States, the participating physicians received the following information in this context [4]:

a) *Prevalence*: The probability that a symptom-free woman aged 52 years has breast cancer (B₊) is 1%.

b) *Sensitivity*: If a symptom-free woman aged 52 years has breast cancer, the probability that she will receive a positive mammography (M₊) result is 80%.

c) *Specificity*: However, if a symptom-free woman aged 52 years does not have breast cancer (B₋), the probability that she will still receive a positive mammography result is 10%.

Based on this information, 95 of the 100 physicians interviewed concluded that – after a positive screening result – the probability of having breast cancer is between 70% and 80%. However, inserting the three given values into Bayes' theorem (see left side of Figure 1) reveals that the PPV is 7.5%! The difficulty of calculating the correct PPVs by means of the given information has been shown in several studies that included physicians [5], medical students [6] and laypeople [7] (see [8] for an overview).

In individual cases, misjudgements may lead to serious errors in decision-making regarding further diagnostics and therapy. Such misjudgements may be avoided if natural frequencies instead of probabilities are used to communicate relevant information. Natural frequencies are the number of different cases occurring in a representative random sample [7]. Usually, the conditional probabilities presented within text books have been derived from natural frequencies. Conversely, probabilities can be easily (re-)translated into natural frequencies. In a first step, *prevalence* is related to a fictitious number of people (in the following, the number of 1000 is used) to calculate the number of people affected by the disease in the random sample (1% of 1000 equals 10; Figure 1, right hand side). In a second step, the number of affected patients receiving a positive result is determined by means of the *sensitivity* of the test (80% of 10 equals 8).

Finally, the number of positive results in the group of healthy people is identified by means of the *specificity* of the test in a third step (the rate of false alarm of 10% in relation to 990 equals 99). Thus, 107 in 1000 women receive a positive result (8+99), but only 8 of these 107 women actually have breast cancer. The quotient $8/107$ is 7.5%, and thus the PPV already mentioned above as a result of Bayes' theorem. Strictly speaking, natural frequencies can also be viewed as applying this rule to a fictitious basic population.

A number of studies [5], [6], [7] (for further examples see [9]) on communicating relevant information have shown that the application of natural frequencies instead of probabilities results in an about threefold increase (from maybe 15% or 20% to approximately 50%) in the percentage of correct conclusions (conclusions consistent with Bayes' theorem). A teaching unit, in which medical students were instructed on how to translate probabilities into natural frequencies and subsequently how to extract the correct solution from there turned out to be much more effective than the traditional method according to which students were introduced to Bayes' theorem and instructed how to insert the respective probabilities [10]. Many women accept their invitation to mammography screening because they hope for a negative result and thus for 'peace of mind'. Are such expectations justified? Figure 1 shows that 893 negative results are to be expected in our fictitious random sample. Here, two types of negative results need to be distinguished: First, breast cancer is overlooked in 2 of 10 women affected by this type of cancer (B+), and, secondly, 891 of the 990 healthy women (B-) receive a correct negative result (2+891=893). Thus, for a woman who could be 99% sure to be not affected by breast cancer without undergoing mammography screening ($1 - \text{prevalence}$), this probability increased by 0.78% to 99.78% ($=891/893$) after the receipt of a negative result. Representation by means of natural frequencies thus helps people understand that even a negative result cannot be equated with security (99.78% does not equal 100%) and that the gain is only marginal (here: 0.78 percentage points).

Risk reduction

Natural frequencies are not only helpful for interpreting positive test results but also for deciding on the implementation of a certain type of diagnostics or therapy. Should women accept the invitation to mammography screening? Should men have their PSA level determined for early detection of prostate cancer? Should people undergo bypass interventions to reduce the risk of heart failure? What is the benefit in comparison to the risks and disadvantages? The main benefit of such medical interventions is risk reduction, for instance, to die of breast or prostate cancer or to have a heart attack. The question is thus: What are the risks without a diagnostic or therapeutic intervention compared to the risks if the intervention was taken?

Let's have a look at the figures for mammography screening. Without screening, 4 in 200 healthy women aged between 50 and 69 years who have not been diagnosed with breast cancer will die of this disease within this period of twenty years [11]. If all women would undergo mammography screening, one less woman would die of breast cancer in this period. The most common ways to communicate the reduction in risk are as follows:

- Relative risk reduction (RRR) amounts to 25%: In 1 of 4 women (=25%), death by breast cancer can be prevented.
- Absolute risk reduction (ARR) amounts to 0.5%: Only 3 in 200 women instead of 4 in 200 women die from breast cancer; 1 in 200 (=0.5%) women can be saved.
- Number-needed-to-screen (NNS) amounts to 200: A total of 200 women have to undergo mammography screening to find the one woman benefitting in terms of surviving the next ten years. Not only the effectiveness of screening methods but also that of therapeutic interventions may be evaluated this way. For therapeutic interventions, this measure has been coined number-needed-to-treat (NNT) and is determined as follows: ARR equals $1/\text{NNT}$ (corresponding to $1/\text{NNS}$ for screening methods).

Note the large difference between the communicated values (25%, 0.5%, and 200). Further note that these values are based on the same data, which often leads to considerable confusion. Which of the three values is relevant for a woman who has received an invitation to mammography screening? She is 1 in 200, so that undergoing the screening procedure only reduces her individual risk by 0.5% – which of course, also applies to the other 199 women who have received the same invitation. The figure of 25% exclusively refers to the 4 women who would die of breast cancer without screening – and nobody knows at this point in time who the 4 women are. If they were known, the screening procedure would not be necessary.

RRR values are commonly used for communicating diagnostic, therapeutic or preventive measures [12], [13]. Whereas ARR values tend to be low as a rule, RRR values are usually high. The use of RRR values in expert literature, the general press and patient information suggests relatively high benefits, but this measure is irrelevant in individual cases. Most people do not understand this value correctly [9], and its application is particularly questionable when the diagnostic, therapeutic, or preventive measures also involve risks [14]. In such cases, the people seeking advice may have decided not for but against the implementation of the measure if risk reduction had been communicated in a more transparent way. A Swiss survey [15] including 53 women showed that most women highly overestimated the benefit of mammography screening and were hardly aware of the risks (false positive results and overtreatment, that means, treatment of patients who have a type of cancer that will be correctly diagnosed but that would never be detected clinically and hence should better not be treated). After the women had

been informed in a clear and transparent manner on RRR and ARR as well as on the relation of these two values, the spontaneous readiness to participate in mammography screening in this study dropped from 68% to 11%.

The explanatory power of survival rates

Another measure which is often used to quantify the benefit of screening programmes is survival rate, mostly the 5-year survival rate. Related to this, Rudy Giuliani, the former Mayor of New York, hit the headlines in 2007. During his electoral campaign and in the context of his candidacy for the office as the President of the United States, Giuliani compared the benefits of the American health care system with those of the British health care system: "I had prostate cancer five, six years ago. My chance of surviving prostate cancer – and, thank God, I was cured of it – in the United States? 82%. My chance of surviving prostate cancer in England? Only 44% under socialized medicine" [16]. After the comparison of the two figures, Giuliani drew a superficially plausible but nevertheless incorrect conclusion. On closer examination, the difference between the two 5-year survival rates had nothing to do with the nationalisation of the health system but with the fact that a screening programme for prostate cancer was available in the United States but not in Britain.

But the conclusion that the availability of a screening programme would reduce the mortality rate would also be wrong. Screening programmes enable early detection of many cancer diseases, but early diagnosis does by no means imply that death can be postponed.

Let's take, as a fictitious example, triplets who simultaneously develop clinically apparent prostate cancer at the age of 83 years and die of the disease at the age of 86 years. The first of the triplets does not participate in any screening programme and dies three years after the spontaneous diagnosis of the disease. The second of the triplets undergoes a PSA test at the age of 80 years, followed by a biopsy and the diagnosis of prostate cancer. The third of the triplets has the luck to meet a unique person at the age of 20 years who is able to tell him on the basis of the form of his earlobes that he has prostate cancer. The disease will not break out for another 63 years but he will die from it after 66 years. What is the triplet's contribution to study results regarding the 5-year survival rate? The first of the triplets will not survive the spontaneous diagnosis by 5 years, and the second is still alive 5 years after the diagnosis was made by means of the PSA test. The case of the third triplet would enhance the reputation of earlobe diagnostics, not only with regard to the 5-year survival rate but also with regard to the 50-year survival rate – a measure unknown in clinical practice. The Lead-time bias affects the three diagnostic methods with their different survival rates differentially: In our fictitious example, such statistics make spontaneous diagnosis look like the worst method and earlobe

diagnosis appear to be the best method. But the fact that the respective survival rates do not allow for any statements on mortality is often overlooked. Regardless of when and why the triplets learn about the diagnosis of cancer, each of them dies the same year.

Lead-time bias is further enhanced by overdiagnosis bias. Screening programmes not only bring forward the time of diagnosis but also further the detection of slowly growing types of cancer, which may never metastasise or manifest clinically. 60% to 80% of men are assumed to develop prostate cancer [17]. Most men do not know about their condition and die of other reasons. What would happen with 5-year survival statistics if such comparatively harmless types of cancer could be detected by means of a highly sensitive test? Such overdiagnoses (correct but rather irrelevant and superfluous diagnoses without any life-extending effects) would push up 5-year survival rates and make the method of diagnosis look rather successful, even if the diagnosis does not at all influence the time of death.

Survival statistics are a good measurement tool for comparing effects of cancer therapies in randomised studies. However, such statistics are useless for comparing groups of patients whose disease was diagnosed by different means (early diagnosis vs. symptom-based discovery). "5-year survival rates are artificially inflated by bringing forward the time of diagnosis and by including tumours with a favourable prognosis. In reality, however, this inflation does not necessarily reduce mortality rates. For this reasons, 5-year survival rates are unsuitable for estimating the effect of early diagnoses" ([16], p 4). Most physicians are unaware of these relations. In one of the respective studies, the percentage of physicians who were able to correctly explain lead-time bias and overdiagnosis bias was less than 10% [18].

Sound knowledge of these distortions seems to be indispensable for estimating the benefits of screening programmes. For a woman diagnosed with breast cancer by means of early detection mammography, Welch and Frankl calculated a probability rate of 13% that death by breast cancer will be avoided because of early diagnosis ([19]; this calculation is based on an assumed reduction in mortality of 20%). In view of such a low probability rate, the authors concluded that "Most women with screen-detected breast cancer have not had their life saved by screening. They are instead either diagnosed early (with no effect on their mortality) or overdiagnosed."

Successful risk communication

The above-mentioned prevalence of breast cancer and test parameters for mammography screening were taken from a U.S. publication of 1982 [4]. We would like to explicitly state that both sensitivity and specificity of mammography screening largely depend on the framework conditions under which programmes are carried out. Significantly less diagnoses of breast cancer will be overlooked and considerably less false positive findings

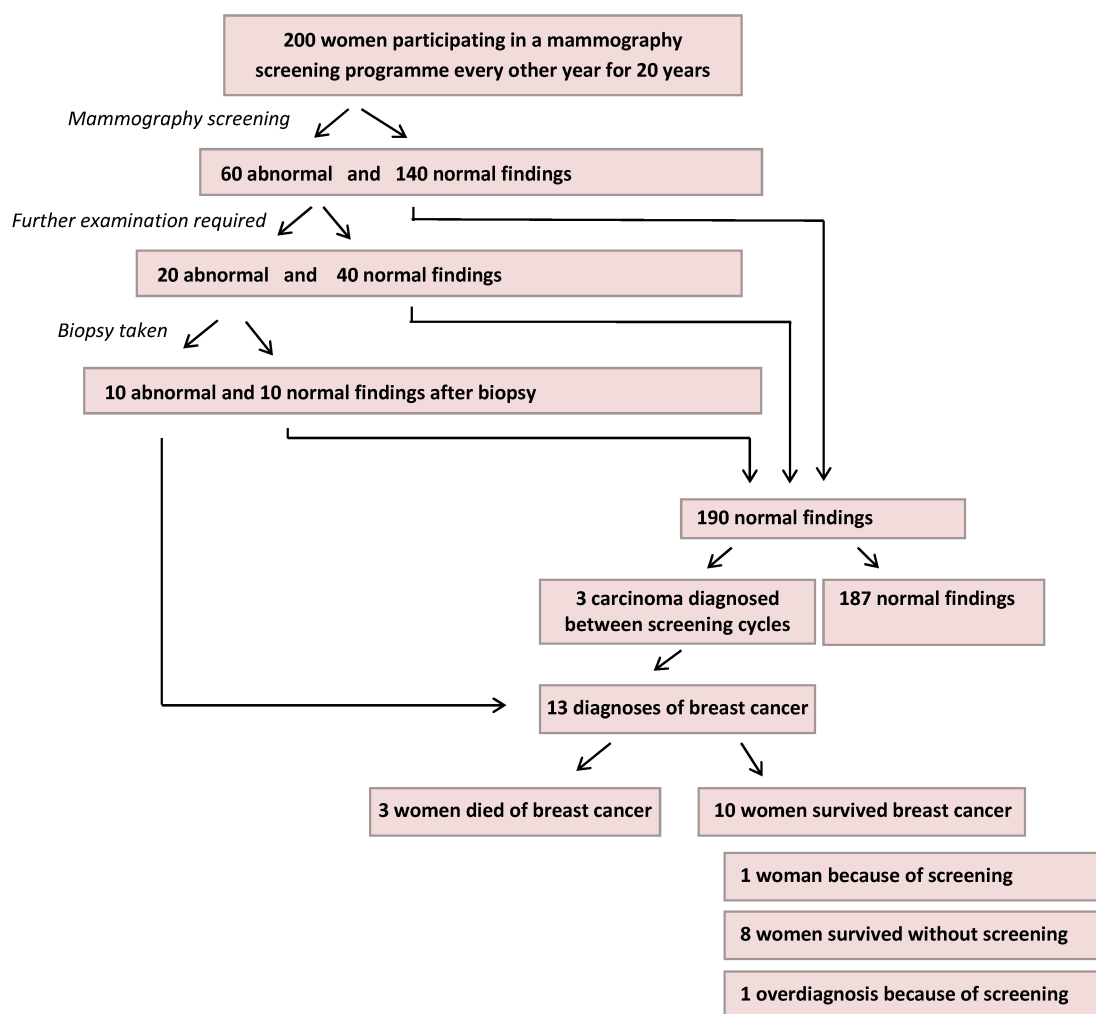


Figure 2: Benefits of mammography screening (adapted from [11])

will occur in quality-assured, systematically conducted screening programmes in which analyses are carried out by specially trained and experienced radiologists than in small gynaecological practices.

Improvements can be observed not only with regard to the figures themselves but also with regard to the manner of their communication. We would like to conclude our article by showing such a positive example. Unfortunately, the number of good examples is rather low (the overview of information material on mammography screening presented in [12] and [13] is rather sobering), but some change is on the way. The presentations designed by the Mammography Cooperative (Kooperationsgemeinschaft Mammographie) in collaboration with the German Cancer Research Centre (Deutsches Krebsforschungszentrum) can be viewed as exemplary, and they are adopted by many physicians and journalists. In these presentations, figures are presented as natural frequencies throughout: the diagnostic properties of the screening (PPV and cases of cancer overlooked) and its benefit (risk reduction) and risks (false alarms and overdiagnoses) relate to one and the same fictitious basic population and are thus directly comparable (see Figure 2). We would like to add that the authors of these presentations are very familiar with the results of studies such as the one mentioned above.

The author put the overview shown in Figure 2 into words ([11], p. 23), which have been published in an information brochure for the general public ([20], p. 10).

The following figures, which are based on experiences made in other countries and on scientific investigations, shall give you a clear idea of how the benefits and risks are roughly distributed within the entire program:

- Of 200 women participating in a mammography screening programme every other year for 20 years, 140 do not receive a suspicious finding. The remaining 60 women require further examination.
- 40 of these 60 women receive a normal finding when further examined, but the remaining 20 women are advised to have a biopsy taken.
- for 10 of these 20 women the suspicion was not confirmed, and the other 10 women receive the diagnosis breast cancer within the screening programme. Over the 20-year period, 3 of the remaining 190 women also receive the diagnosis of breast cancer but between two screening circles.
- 3 of the overall 13 women with the diagnosis of breast cancer die of the disease, and 10 women do not.

- 1 of these 10 women would not have learned about her diagnosis of breast cancer without the mammography screening programme; 8 women would have been successfully treated, even without participating in the screening programme but some of them would have required a more arduous course of treatment. 1 in 200 women is saved from death by breast cancer because of her regular participation in the mammography screening programme.

This overview meets all criteria of the catalogue compiled by the specialist team for patient information of the German Network for Evidence-based Medicine (DNEbM), which was developed to support physicians in counselling patients on early cancer diagnosis [21], [22]. The overview also corresponds with the demands for better risk communication in the context of screening programmes [23]. The overview is transparent, and the manner of communication of the most important figures allows for a direct comparison of benefits and risks. This way, every woman is able to decide, either by herself or after consultation with her physician, if she would like to participate in the lottery – also termed mammography screening (further commendable presentations are available in [24], [25], [26]. We would like to encourage physicians, expert societies, patient organisations, health insurances and authorities to take up this example and compile further transparent overviews on the diagnosis and treatment of diseases. The Harding Centre for Risk Literacy at the Max-Planck-Institute for Human Development in Berlin refers to such overviews with the term ‘fact box’ (see also [27] and [28]) and has already produced a number of these boxes [29]. Risks are unavoidable. They have always been and will always be around – but poor risk communication and misunderstanding are really unnecessary.

Notes

Competing interests

The authors declare that they have no competing interests.

References

- Ahmed H, Naik G, Willoughby H, Edwards AG. Communicating risk. *BMJ*. 2012;344:e3996. DOI: 10.1136/bmj.e3996
- Gigerenzer G, Muir Gray JA, editors. Better doctors, better patients, better decisions: Envisioning health care 2020. Cambridge: MIT; 2011.
- Malek D, Käab-Sanyal V; Kooperationsgemeinschaft Mammographie. Evaluationsbericht 2011 – Zusammenfassung der Ergebnisse des Mammographie-Screening-Programms in Deutschland. Berlin: Kooperationsgemeinschaft Mammographie; 2014. Available from: http://www.mammo-programm.de/download/MAMMO_EvalBericht_20141201_web%282%29.pdf [cited 2015 Mar 8]
- Eddy DM. Probabilistic reasoning in clinical medicine: Problems and opportunities. In: Kahneman D, Slovic P, Tversky A, editors. *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press; 1982. pp. 249-67. DOI: 10.1017/CBO9780511809477.019
- Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Acad Med*. 1998 May;73(5):538-40. DOI: 10.1097/00001888-199805000-00024
- Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G. Medicine. Communicating statistical information. *Science*. 2000 Dec;290(5500):2261-2. DOI: 10.1126/science.290.5500.2261
- Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychol Rev*. 1995;102:684-704. DOI: 10.1037/0033-295X.102.4.684
- Ghosh AK, Ghosh K. Translating evidence-based information into effective risk communication: current challenges and opportunities. *J Lab Clin Med*. 2005 Apr;145(4):171-80. DOI: 10.1016/j.lab.2005.02.006
- Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients to make sense of health statistics. *Psychol Sci Public Interest*. 2007;8:53-96. DOI: 10.1111/j.1539-6053.2008.00033.x
- Kurzenhäuser S, Hoffrage U. Teaching Bayesian reasoning: an evaluation of a classroom tutorial for medical students. *Med Teach*. 2002 Sep;24(5):516-21. DOI: 10.1080/0142159021000012540
- Weymayr C; Kooperationsgemeinschaft Mammographie. Kennzahlen Mammographie-Screening – Dokumentation 2010 Version 1.2. Berlin: Kooperationsgemeinschaft Mammographie; 2010. Available from: <http://www.komen.de/html/img/pool/1kennzahlenmammographie-screeningdokumentationv1.2.pdf> [cited 2015 Mar 8]
- Kurzenhäuser S. Welche Informationen vermitteln deutsche Gesundheitsbroschüren über die Screening-Mammographie [What kind of information do German health information pamphlets provide on mammography screening?]. *Z Arztl Fortbild Qualitätssich*. 2003 Feb;97(1):53-7.
- Kurzenhäuser S, Hoffrage U. Designing risk communication in health. In: Todd PM, Gigerenzer G, ABC Research Group, editors. *Ecological rationality – intelligence in the world*. New York: Oxford University Press; 2012. p. 428-53.
- Woloshin S, Schwartz LM. Numbers needed to decide. *J Natl Cancer Inst*. 2009 Sep;101(17):1163-5. DOI: 10.1093/jnci/djp263
- Matter-Walstra K, Hoffrage U. Individuelle Entscheidungsfindung am Beispiel der Brustkrebs-Früherkennung—Erfahrungen aus Fokusgruppen in der Schweiz. [Individual decision making concerning breast cancer screening—Observations with focus groups in Switzerland]. *Schweizer Zeitschr Managed Care Care Management*. 2001;3:26-9.
- Wegwarth O, Gigerenzer G. Risikokommunikation – Risiken und Unsicherheiten richtig verstehen lernen. *Dtsch Arztebl*. 2011;108(9):A-448-51.
- Sökeland J, Weiss HE. Häufigkeit des Prostatakarzinoms. 2011 [last update 2014 Jul 28]. Available from: http://www.prostata.de/pca_haeufigkeit.html [cited 2015 Mar 8]
- Wegwarth O, Gaissmaier W, Gigerenzer G. Deceiving numbers: survival rates and their impact on doctors' risk communication. *Med Decis Making*. 2011 May-Jun;31(3):386-94. DOI: 10.1177/0272989X10391469
- Welch HG, Frankel BA. Likelihood that a woman with screen-detected breast cancer has had her "life saved" by that screening. *Arch Intern Med*. 2011 Dec;171(22):2043-6. DOI: 10.1001/archinternmed.2011.476

20. Gemeinsamer Bundesausschuss. Informationen zum Mammographiescreening. Available from: http://www.mammo-programm.de/download/merkblatt_deutsch_web.pdf [cited 2015 Mar 8]
21. Koch K, Mühlhauser I. Kriterien zur Erstellung von Patienteninformationen zu Krebsfrüherkennungsuntersuchungen: Stellungnahme des Fachbereichs Patienteninformation des Deutschen Netzwerkes für Evidenzbasierte Medizin (DNEbM). 2008. Available from: <http://www.ebm-netzwerk.de/pdf/stellungnahmen/dnebm-080630.pdf> [cited 2015 Mar 8]
22. Griebenow B. Beratung zur Krebsfrüherkennung: Vor- und Nachteile darstellen. Dtsch Ärztebl. 2008;05(33):A-1724.
23. Jørgensen KJ, Brodersen J, Hartling OJ, Nielsen M, Gøtzsche PC. Informed choice requires information about both benefits and harms. J Med Ethics. 2009 Apr;35(4):268-9. DOI: 10.1136/jme.2008.027961
24. Leitlinienprogramm Onkologie der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V., der Deutschen Krebsgesellschaft e. V. und der Deutschen Krebshilfe e. V. Früherkennung von Brustkrebs – Eine Entscheidungshilfe für Frauen. 2010. Available from: http://leitlinienprogramm-onkologie.de/uploads/tx_sbdownloader/Patientenleitlinie_Brustkrebs_Fruherkennung.pdf [cited 2014 Nov 25]
25. Medizinische Universität Graz, EBM Review Center. Mammographie basierte Brustkrebsfrüherkennung – Recherche und Aufbereitung von Kennzahlen für eine informierte Entscheidung. Available from: http://www.frueh-erkennen.at/tl_files/bkfp/Dokumente/Fuer%20Radiologen/Mammographie%20basierte%20Screening-Kennzahlen%20informierte%20Entscheidung.pdf [cited 2015 Mar 8]
26. Gøtzsche PC, Hartling OJ, Nielsen M, Brodersen J. Screening für Brustkrebs mit Mammographie. Copenhagen: Nordisches Cochrane Zentrum; 2012. Available from: <http://www.cochrane.dk/screening/mammografi-de.pdf> [cited 2015 Mar 8]
27. Schwartz LM, Woloshin S, Welch HG. The drug facts box: providing consumers with simple tabular data on drug benefit and harm. Med Decis Making. 2007 Sep-Oct;27(5):655-62. DOI: 10.1177/0272989X07306786
28. Schwartz LM, Woloshin S, Welch HG. Using a drug facts box to communicate drug benefits and harms: two randomized trials. Ann Intern Med. 2009 Apr;150(8):516-27. DOI: 10.7326/0003-4819-150-8-200904210-00106
29. Harding Zentrum für Risikokompetenz. Faktenboxen. Available from: <https://www.harding-center.mpg.de/de/gesundheitsinformationen/faktenboxen> [cited 2015 Mar 8]
30. Hoffrage U, Kurzenhäuser S, Gigerenzer G. Wie kann man die Bedeutung medizinischer Testbefunde besser verstehen und kommunizieren [How can one improve the understanding and communication of the importance of medical test results?]. Z Arztl Fortbild Qualitatssich. 2000 Oct;94(9):713-9.

Corresponding author:

Prof. Dr. Ulrich Hoffrage
Faculty of Business and Economics, University of
Lausanne, 1015 Lausanne, Switzerland, Phone: +41 21
692 3490, Fax: +41 21 692 3305
Ulrich.hoffrage@unil.ch

Please cite as

Hoffrage U, Koller M. Chances and risks in medical risk communication. GMS Ger Med Sci. 2015;13:Doc07.
DOI: 10.3205/000211, URN: urn:nbn:de:0183-0002117

This article is freely available from

<http://www.egms.de/en/journals/gms/2015-13/000211.shtml>

Received: 2014-12-02

Revised: 2015-05-04

Published: 2015-07-06

Copyright

©2015 Hoffrage et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

Chancen und Risiken der Risikokommunikation in der Medizin

Zusammenfassung

Kommunikation, die im Alltag zwischen Ärzten und Patienten stattfindet, wird durch eine Reihe störender Faktoren begleitet. Neben spezifischen Interessen sind es vor allem Fehler und Missverständnisse, die von beiden Seiten zum Risiko der Risikokommunikation beitragen. Am Beispiel des Mammographie-Screenings wird dargestellt, wie die Bedeutung von Testergebnissen und die Aussagekraft von Maßnahmen, die zur Reduktion von Risiken ergriffen werden, oft missverstanden wird. Ferner wird gezeigt, wie eine gelungene Risikokommunikation aussehen kann.

Schlüsselwörter: Risikokommunikation, Mammographie-Screening, Bayes'sche Regel, Wahrscheinlichkeiten, natürliche Häufigkeiten, relative Risikoreduktion, Überlebensraten

Patienten und Ärzte sitzen in einem Boot

„Zu Risiken und Nebenwirkungen fragen Sie Ihren Arzt oder Apotheker“ – wer kennt ihn nicht, diesen Ratschlag. Genaugenommen ist es allerdings nicht nur ein gutgemeinter Ratschlag für den Leser einer Medikamentenwerbung, sondern auch ein Passus, mit dem sich ein Pharmaunternehmen gegen etwaige Regressforderungen schützt. Dieses Beispiel zeigt: Risiken und Nebenwirkungen sind etwas Schlechtes, jeder will sie vermeiden und Risikokommunikation kann vielschichtig sein.

Auf der einen Seite wollen Bürger vermeiden zu Patienten zu werden und Patienten wollen Nebenwirkungen von Medikamenten vermeiden; auf der anderen Seite wollen Ärzte und Pharmaunternehmen vermeiden, Fehler zu begehen, missverstanden zu werden, oder verklagt zu werden. Mit dem Bestreben, Risiken zu vermeiden, sind die Gemeinsamkeiten noch lange nicht erschöpft. So ist jede der beiden Seiten auf die jeweils andere angewiesen: Während Patienten die Hilfe von Experten brauchen, sind die diversen Akteure im Gesundheitswesen nicht nur für, sondern auch durch die Patienten da, denn ohne diese Patienten hätten sie keine Aufgaben, ja es gäbe ihren Beruf gar nicht. Außerdem möchte jede der beiden Seiten der jeweils anderen vertrauen: Der leidende Patient gewährt dem herannahenden Arzt einen Vertrauensvorschluss, zumal dieser in der Regel auch die erhoffte Hilfe bringt; und der Arzt darf darauf vertrauen, dass sich der geheilte Patient dankbar erweist.

Nicht immer geht es dem Patienten nach der Behandlung besser, aber auch in diesem Fall setzen sich die Gemein-

Ulrich Hoffrage¹

Michael Koller²

1 Faculty of Business and Economics, University of Lausanne, Lausanne, Switzerland

2 Centre for Clinical Studies, University Hospital Regensburg, Regensburg, Germany

samkeiten fort: Ärzte machen manchmal Fehler, und manchmal handeln sie nicht ausschließlich zum Wohle des Patienten, weil sie selber Zwängen unterliegen oder einem Interessenkonflikt ausgesetzt sind. Da dies vielen Patienten durchaus bewusst ist, mischt sich in das Vertrauen, das sie dem Arzt entgegenbringen, manchmal auch eine Prise Misstrauen – oft auch noch geschürt durch entsprechende Kommunikation: Zu Risiken und Nebenwirkungen im Umgang mit Ärzten fragen Patienten oft andere Patienten. Aber auch Patienten machen manchmal Fehler. Sie halten sich nicht an die Anweisungen des Arztes oder sie machen den Arzt zu Unrecht für eine Verschlechterung ihres Zustandes verantwortlich – was gelegentlich auch zu rechtlichen Auseinandersetzungen führt. Entsprechend wissen auch Ärzte so manches von den Risiken und Nebenwirkungen im Umgang mit Patienten zu berichten.

Von den Risiken der Risikokommunikation

Die aufgeführten Gemeinsamkeiten zwischen Ärzten und Patienten, die Möglichkeiten sich zu helfen, aber auch sich gegenseitig einen Schaden zuzufügen, lassen sich auch im Bereich der Risikokommunikation feststellen. Laien können oft nicht beurteilen, wie gefährlich beispielsweise ihre Beschwerden sind und was man dagegen tun kann. Experten können hier helfen und die relevanten Informationen zur Verfügung stellen. Genau darin liegen die Chancen der Risikokommunikation [1].

Aber – um Benjamin Franklin zu zitieren – „*nichts ist sicher außer dem Tod und der Steuer*“. So können auch die Auskünfte von Experten fehlerhaft sein. Oft sind sie

auch lediglich unvollständig oder missverständlich, wobei es dann nicht immer klar ist, ob eher der Arzt oder eher der Patient für das Missverständnis verantwortlich ist. In diesen Fehlern und möglichen Missverständnissen liegen die Risiken der Risikokommunikation.

Idealerweise führt das Wissensgefälle zwischen Arzt und Patient zu einer transparenten Vermittlung relevanter Informationen. Der Alltag sieht aber oft anders aus, die hier stattfindende Kommunikation über Risiken ist nicht frei von störenden Interessen [2]. Zum einen kann ein medizinischer Betrieb nur aufrechterhalten werden, wenn seine Kosten gedeckt sind. Zum anderen ist es verständlich und legitim, wenn sich Ärzte und Pflegepersonal, aber auch Behörden oder Klinikmanagement vor rechtlichen Auseinandersetzungen schützen wollen. Darüber hinaus spielen oft auch institutionelle Gegebenheiten in die Art und Weise, wie Risiken kommuniziert werden, mit hinein. Das vielleicht gravierendste Hindernis für eine transparente und verständliche Risikokommunikation liegt, unserer Auffassung nach, jedoch in der Unfähigkeit mancher Beratenden, mit Zahlen richtig umzugehen, das heißt, diese richtig zu verstehen und richtig zu vermitteln. Im Folgenden werden wir für drei Bereiche aufzeigen, wie bestimmte Werte und Statistiken, die häufig bei der Kommunikation von Risiken verwendet werden, missverstanden werden können – und wie sie kommuniziert werden können, um das Risiko von Missverständnissen zu reduzieren.

Bedeutung von Testergebnissen (Bayesianische Inferenzen)

Eine 52-jährige Frau ist der Einladung zu einem Mammographie-Screening gefolgt und hat nun, obwohl sie keine einschlägigen Symptome aufwies, plötzlich einen „auffälligen“ Befund erhalten, d.h. es wurde eine verdächtige Struktur entdeckt und es muss näher abgeklärt werden, ob sich dahinter wirklich ein Mammakarzinom verbirgt. Im Jahre 2011 wurden in Deutschland ca. 4,9 Millionen Frauen im Alter von 50–69 zu dieser Reihenuntersuchung eingeladen. Von den ca. 2,7 Millionen Frauen, die dieser Einladung gefolgt sind, wurden 8,6% (unter den ca. 800.000 Erstuntersuchungen) bzw. 3,2% (unter den ca. 1,9 Millionen Folgeuntersuchungen) erneut einbestellt [3]. Was bedeutet aber eine derartige Wiedereinbestellung genau? Der häufigste Grund dafür ist ein positiver Befund bei der ersten Mammographie. Wie hoch ist die Wahrscheinlichkeit, dass in einem solchen Fall tatsächlich Brustkrebs vorliegt? Diese Wahrscheinlichkeit, der sogenannte *positive prädiktive Wert* (PPW), hängt von drei Parametern ab. In einer in den USA durchgeführten Untersuchung wurden den teilnehmenden Ärzten dazu folgende Informationen gegeben [4]:

- a) *Prävalenz*: Die Wahrscheinlichkeit, dass eine symptomfreie Frau im Alter von 52 Jahren Brustkrebs (B+) hat, beträgt 1%.
- b) *Sensitivität*: Wenn eine dieser Frauen Brustkrebs hat, dann beträgt die Wahrscheinlichkeit, dass sie

einen positiven Mammographie-Befund (M+) erhält, 80%.

c) *Spezifität*: Wenn eine dieser Frauen jedoch keinen Brustkrebs (B-) hat, dann beträgt die Wahrscheinlichkeit, dass sie dennoch einen positiven Mammographie-Befund erhält, 10%.

95 von 100 der befragten Ärzte schlossen aus diesen Informationen, dass die Wahrscheinlichkeit für Brustkrebs nach einem positiven Testergebnis zwischen 70 und 80 Prozent liegen würde. Setzt man die gegebenen Werte jedoch in die so genannte Bayes'sche Regel ein (linke Seite der Abbildung 1), ergibt sich ein PPW von 7,5 Prozent! Diese Schwierigkeiten, aus den gegebenen Wahrscheinlichkeiten den richtigen PPW abzuleiten, konnten in diversen Studien mit Ärzten [5], Medizinstudenten [6] und medizinischen Laien [7] vielfach repliziert werden (für einen Überblick siehe [8]).

Diese Fehleinschätzungen, die im Einzelfall zu gravierenden Fehlentscheidungen bezüglich weiterer Diagnostik und Therapie führen können, lassen sich vermeiden, wenn man bei der Kommunikation der relevanten Informationen nicht Wahrscheinlichkeiten, sondern natürliche Häufigkeiten verwendet. Natürliche Häufigkeiten sind die Anzahlen der verschiedenen Fälle, die in einer repräsentativen Stichprobe auftreten [7]. In der Regel sind es diese Häufigkeiten, die den heute üblicherweise in Lehrbüchern angegebenen bedingten Wahrscheinlichkeiten zugrunde liegen.

Wahrscheinlichkeiten lassen sich leicht in natürliche Häufigkeiten (zurück) übersetzen. Im ersten Schritt wird dabei die *Prävalenz* auf eine fiktive Anzahl von Personen bezogen (im Folgenden: 1000), um so die Anzahl der Erkrankten in dieser Stichprobe zu ermitteln (1% von 1000 ergibt 10; Abbildung 1, rechts). Im zweiten Schritt wird die *Sensitivität* des Tests benutzt um zu ermitteln, wie viele der Erkrankten ein positives Ergebnis erhalten (80% von 10 ergibt 8). Und im dritten Schritt wird mit Hilfe der *Spezifität* des Tests festgestellt, wie viele positive Befunde es in der Gruppe der Gesunden gibt (die Falsch-Alarm Rate von 10%, bezogen auf 990, ergibt 99). Damit lässt sich ersehen, dass von 1000 Frauen 107 einen positiven Befund erhalten (8+99), zugleich aber von diesen 107 nur 8 tatsächlich einen Brustkrebs haben. Der Quotient $8/107$ ergibt 7,5% – den PPW, den wir oben bereits als Ergebnis der Bayes'schen Regel genannt hatten. Genau genommen können also natürliche Häufigkeiten als eine Anwendung dieser Regel auf eine fiktive Grundgesamtheit gesehen werden.

Eine Reihe von Studien [5], [6], [7] (für weitere Beispiele siehe [9]) hat gezeigt, dass die Verwendung von natürlichen Häufigkeiten (statt Wahrscheinlichkeiten) bei der Kommunikation von relevanten Information den Prozentsatz der richtigen Schlüsse (d.h., solche die konsistent mit der Bayes'schen Regel sind) um rund das dreifache ansteigen lässt (von 15–20% auf etwa 50%). Des Weiteren war diese Methode in einer Lehreinheit für Medizinstudenten weit effektiver als die traditionelle Methode, nach der den Studenten die Bayes'sche Regel beige-

Wahrscheinlichkeiten

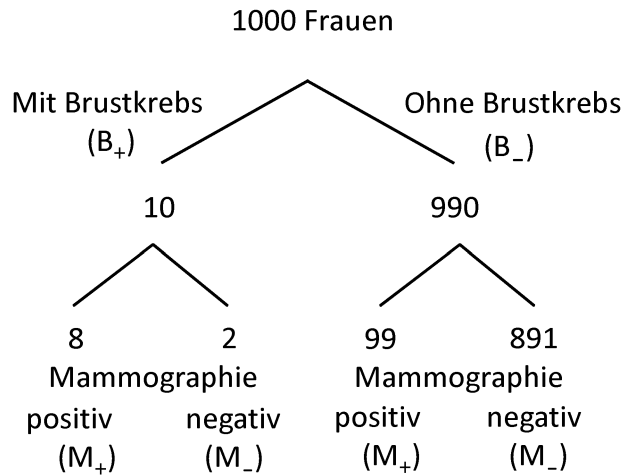
$$p(B_+) = 0,01$$

$$p(M_+ | B_+) = 0,80$$

$$p(M_+ | B_-) = 0,10$$

$$p(B_+ | M_+) = \frac{0,01 \cdot 0,80}{0,01 \cdot 0,80 + 0,99 \cdot 0,10} = 0,075$$

Natürliche Häufigkeiten



$$p(B_+ | M_+) = \frac{8}{8 + 99} = 0,075$$

Abbildung 1: Repräsentation ein und derselben Information in Wahrscheinlichkeiten und natürlichen Häufigkeiten (adaptiert aus [30]).

bracht wurde, und ihnen gezeigt wurde, wie die entsprechenden Wahrscheinlichkeiten einzusetzen sind [10]. Viele Frauen nehmen deshalb am Brustkrebscreening teil, weil sie sich ein negatives Ergebnis erhoffen und dadurch dann endlich „Sicherheit“ haben. Ist diese Erwartung berechtigt? Aus Abbildung 1 ist leicht zu ersehen, dass in unserer fiktiven Stichprobe 893 negative Ergebnisse zu erwarten sind. Dabei sind zwei Fälle zu unterscheiden: Zum einen wird bei 2 von den 10 erkrankten Frauen (B+) der Brustkrebs übersehen und zum anderen erhalten von den 990 gesunden Frauen (B-) 891 einen richtig-negativen Befund (2+891=893). Für eine Frau, die ohne Mammographie eine 99%ige Sicherheit haben konnte, keinen Brustkrebs zu haben (1 – Prävalenz), ist diese Wahrscheinlichkeit nach Erhalt eines negativen Ergebnisses also um 0,78% auf 99,78% (=891/893) gestiegen. Eine Darstellung mit natürlichen Häufigkeiten hilft also zu verstehen, dass auch ein negatives Ergebnis noch keine Sicherheit bedeutet (99,78% sind immer noch nicht 100%) und dass der Zugewinn oft nur gering ist (hier: 0,78 Prozentpunkte).

Reduktion von Risiken

Natürliche Häufigkeiten sind nicht nur hilfreich, wenn es um die Interpretation von positiven Testbefunden geht, sondern auch wenn es darum geht zu entscheiden, ob eine bestimmte Diagnostik oder Therapie einzusetzen ist. Soll eine Frau der Einladung zum Brustkrebs-Screening folgen? Soll ein Mann seinen PSA-Wert bestimmen lassen, um ggf. Prostatakrebs rechtzeitig zu entdecken? Soll ein Bypass gelegt werden, um das Risiko eines

Herzinfarkts zu senken? Wie steht hier jeweils der Nutzen im Verhältnis zu den Risiken und Nachteilen? Der hauptsächliche Nutzen derartiger Maßnahmen liegt in einer Verminderung des Risikos, beispielsweise an Brustkrebs oder an Prostatakrebs zu sterben, oder einen Herzinfarkt zu erleiden. Die Frage lautet daher: Wie verhalten sich die Risiken ohne Diagnostik zu den Risiken, die mit den diagnostischen und therapeutischen Maßnahmen verbunden sind?

Betrachten wir als Beispiel hier die Zahlen für das Brustkrebscreening. Ohne Screening würden 4 von 200 gesunden Frauen im Alter von 50 bis 69 Jahren, bei denen noch kein Brustkrebs diagnostiziert wurde, in einem Zeitraum von 20 Jahren an Brustkrebs sterben [11]. Würden alle am Brustkrebscreening teilnehmen, würde eine weniger in diesem Zeitraum an Brustkrebs sterben. Die gängigsten Maße, um diese Reduktion des Risikos zu kommunizieren, sind:

- Die relative Risikoreduktion (RRR). Diese beträgt 25%: Bei 1 von 4 (=25%) kann ein Tod durch Brustkrebs verhindert werden.
- Die absolute Risikoreduktion (ARR). Diese beträgt 0,5%: Statt 4 von 200 sterben nur noch 3 von 200; 1 von 200 (=0,5%) kann gerettet werden.
- Die Number-needed-to-screen (NNS). Diese beträgt 200: Es müssen 200 Frauen am Screening teilnehmen, um die eine zu finden, die im Sinne eines verlängerten Lebens profitiert. Nicht nur die Effizienz von Screeningverfahren, sondern auch die von therapeutischen Verfahren lässt sich mit diesem Maß bewerten; dann spricht man von number-needed-to-treat (NNT). Es gilt $ARR=1/NNS$ (bzw. $1/NNT$).

Man beachte, dass die kommunizierten Werte sehr unterschiedlich sind (25%, 0,5%, 200). Dennoch beruhen alle auf denselben Daten. Entsprechend groß ist oft die Verwirrung. Welches der drei Werte ist nun für die Frau relevant, die die Einladung zum Screening bekommen hat? Sie ist eine von 200, und für sie kann deshalb die Teilnahme auch nur eine Reduzierung ihres individuellen Risikos um 0,5% mit sich bringen – gleiches gilt natürlich auch für die anderen 199, die den Einladungsbrief in den Händen halten. Die 25% hingegen beziehen sich ausschließlich auf genau die 4 Frauen, die ohne Screening an Brustkrebs sterben würden – und niemand weiß zu diesem Zeitpunkt, welche 4 es sein werden. Wüsste man es, bräuchte man das Screening nicht.

Üblicherweise wird die RRR verwendet, wenn der Nutzen von diagnostischen, therapeutischen oder prophylaktischen Maßnahmen kommuniziert wird [12] (siehe auch [13]). Während die ARR in der Regel niedrig ist, ist die RRR in der Regel hoch. Somit wird durch die weite Verbreitung des RRR sowohl in der Fachliteratur als auch in der allgemeinen Presse und in Patientenbroschüren ein Maß angegeben, das einen relativ hohen Nutzen suggeriert, welcher aber für den Einzelnen irrelevant ist. Dieses Maß wird von den meisten missverstanden [9]. Seine Verwendung ist dann besonders fraglich, wenn mit der Maßnahme auch Nachteile verbunden sind [14]. Dabei kann es durchaus vorkommen, dass sich der Ratsuchende nicht für, sondern gegen die Durchführung der Maßnahme entschieden hätte, wenn ihm die Risikoreduktion transparenter kommuniziert worden wäre. In einer Untersuchung mit 53 Frauen in der Schweiz [15] kam zu Tage, dass die meisten Frauen den Nutzen enorm überschätzten und die Nachteile (falsch-positive Befunde und Überbehandlungen, d.h. Behandlung von richtig diagnostizierten Krebsfällen, die nie klinisch auffällig geworden wären) kaum kannten. In dieser Studie fiel die spontane Teilnahmebereitschaft von 68% drastisch auf 11% ab, nachdem den Frauen sowohl die RRR als auch die ARR und ihr Verhältnis zueinander transparent erklärt worden sind.

Aussagekraft von Überlebensraten

Ein anderes Maß, das oft verwendet wird, um den Nutzen von Screening-Programmen zu quantifizieren und zu belegen, ist die Überlebensrate, wobei meist die Fünfjahresüberlebensrate genommen wird. Für ein kurioses, historisches Aperçu sorgte Rudy Giuliani, ehemaliger Bürgermeister von New York, im Jahre 2007. Während eines Wahlkampfauftritts im Rahmen seiner Kandidatur um das Amt des Präsidenten der USA stellte er 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“ (deutsche Übersetzung entnommen [16]).

Giuliani zog aus dem Vergleich der beiden Zahlen einen zwar vordergründig plausiblen, aber dennoch falschen Schluss. Bei näherer Betrachtung hat der Unterschied zwischen den Fünfjahresüberlebensraten nichts mit der Verstaatlichung des Gesundheitswesens zu tun, sondern damit, dass es in den USA ein Prostatafrüherkennungsprogramm gab, in Großbritannien aber nicht.

Aber auch der Schluss, dass ein Früherkennungsprogramm die Sterblichkeit senken würde, wäre falsch. Durch ein Screening-Programm werden viele Krebserkrankungen früher entdeckt, aber eine derartige Vorverlegung des Diagnosezeitpunkts heißt noch lange nicht, dass dadurch auch der Sterbezeitpunkt hinausgeschoben werden kann. Nehmen wir als ein fiktives Beispiel Drillinge, die im Alter von 83 Jahren gleichzeitig einen klinisch auffälligen Prostatakrebs entwickeln und im Alter von 86 Jahren daran versterben. Der erste nimmt an keinem Früherkennungsprogramm teil und stirbt drei Jahre nach der spontanen Entdeckung der Erkrankung daran. Der zweite lässt im Alter von 80 Jahren einen PSA-Test durchführen, der zu einer Biopsie und dann schließlich zur Diagnose Krebs führt. Und der dritte hat das seltene Glück, im Alter von 20 Jahren einen einzigartigen Menschen kennenzulernen, der ihm aufgrund der Form seines Ohrläppchens sagen kann, dass er einen Prostatakrebs hat, der noch 63 Jahre schlummern und dann nach 66 Jahren zum Tode führen wird. Welchen Beitrag werden die Drillinge zu Studien liefern, in den Fünfjahresüberlebensraten ermittelt werden? Der erste wird seine Diagnose, die durch spontane Entdeckung zustande kam, nicht um fünf Jahre überleben; der zweite ist fünf Jahre nach der durch den PSA Test ausgelösten Diagnose noch am Leben; und der dritte würde sogar dazu beitragen, die Ohrläppchendiagnostik nicht nur bezüglich ihrer Fünfjahresüberlebensrate, sondern auch bezüglich ihrer Fünfzigjahresüberlebensrate – ein Maß, welches in der Praxis nicht verwendet wird – gut dastehen zu lassen. Der Vorlauf-Bias (*leadtime bias*) wirkt sich auf die unterschiedlichen Diagnosemethoden mit ihren jeweiligen Fünfjahresüberlebensraten unterschiedlich aus. In unserem fiktiven Beispiel lassen diese Statistiken die spontane Entdeckung als am schlechtesten und die Ohrläppchenmethode als die beste dastehen. Aber dabei wird gerne übersehen, dass die jeweiligen Überlebensraten keine Aussagen über die Sterblichkeit als solche erlauben. Ungeachtet der Tatsache, wann und wodurch die Drillinge von ihrem Krebs erfahren, versterben sie alle im gleichen Jahr.

Der Vorlauf-Bias wird durch den Überdiagnose-Bias zusätzlich verstärkt. Durch Screeningprogramme wird nicht nur der Diagnosezeitpunkt vorverlegt, sondern es werden auch solche Arten von Krebs entdeckt, die langsam wachsen, ggf. nie metastasieren würden, und möglicherweise sogar nie klinisch auffällig werden würden. Beim Prostatakrebs wird angenommen, dass bis zu 60% der über 80-jährigen Männer daran erkrankt sind [17]. Die meisten wissen es nicht und sterben aus anderen Gründen. Was würde mit den Fünfjahresüberlebensstatistiken passieren, wenn ein sehr sensitiver Test auch solche vergleichsweise harmlosen Krebsarten entdecken würde?

Derartige Überdiagnosen (richtige, aber eigentlich irrelevante und überflüssige Diagnosen, die keine lebensverlängernden Effekte haben) würden die Fünfjahresüberlebensraten in die Höhe treiben und die Diagnosemethode entsprechend gut dastehen lassen – und das würden sie selbst dann tun, wenn die Diagnose nicht den geringsten Einfluss auf den Todeszeitpunkt hätte.

Überlebensstatistiken sind ein gutes Messinstrument für Vergleiche von Effekten der Krebstherapien in randomisierten Studien, aber sie taugen überhaupt nicht für Vergleiche von Gruppen, deren Erkrankung unterschiedlich diagnostiziert wurde (durch Früherkennung versus symptom-basierte Entdeckung). „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“ ([16], S. 4). Diese Zusammenhänge sind den meisten Ärzten nicht bekannt; in einer dieser Studien lag der Anteil der Ärzte, die den Vorlauf-bias und den Überdiagnose-bias korrekt erklären konnten, bei unter 10% [18].

Und dabei scheint gerade das Wissen um diese beiden Verzerrungen unerlässlich zu sein wenn es darum geht, den Nutzen eines Screening abschätzen zu können. So berechnen Welch und Frankl für eine Frau, bei der durch eine Früherkennungsmammographie ein Brustkrebs entdeckt wird, eine Wahrscheinlichkeit von 13%, dass durch genau diese frühere Diagnose ein Tod durch Brustkrebs verhindert wird ([19]; diese Berechnung beruht auf einer angenommenen Mortalitätsreduktion von 20%). Angesichts dieser niedrigen Wahrscheinlichkeit kommen die Autoren zu dem Schluss, dass „Most women with screen-detected breast cancer have not had their life saved by screening. They are instead either diagnosed early (with no effect on their mortality) or overdiagnosed.“

Gelungene Risikokommunikation

Die oben angegebene Prävalenz von Brustkrebs sowie Testparameter des Mammographie-Screenings sind entnommen einer amerikanischen Publikation aus dem Jahre 1982 [4]. Es sei hier ausdrücklich betont, dass sowohl die Sensitivität und auch die Spezifität der Mammographie stark von den Rahmenbedingungen abhängen, unter denen diese durchgeführt wird. In einem qualitätsgesicherten, systematisch durchgeführten Screening mit vielen Befundungen von speziell ausgebildeten und erfahrenen Radiologen wird es deutlich weniger übersehene Brustkrebs- und weniger falsch-positive Befunde geben als in kleinen gynäkologischen Praxen.

Nicht nur bezüglich der Zahlen, sondern auch bezüglich der Art und Weise, wie man sie kommuniziert, gibt es Verbesserungen. Genau mit einem solchen positiven Beispiel wollen wir diesen Aufsatz beschließen. Leider gibt es deren nicht allzu viele (für einen recht ernüchter-

den Überblick über Informationsmaterial zum Mammographie-Screening siehe [12], [13]), aber es scheint sich hier etwas zu verändern. So können die Darstellungen der Kooperationsgemeinschaft Mammographie, die in Zusammenarbeit mit dem Deutschen Krebsforschungszentrum erstellt wurde, als vorbildlich betrachtet werden (und diese werden von vielen Ärzten und Journalisten übernommen). Bei der Darstellung des Zahlenmaterials werden durchgängig natürliche Häufigkeiten verwendet: sowohl die diagnostischen Eigenschaften des Screenings (PPW und übersehene Fälle von Krebs), dessen Nutzen (Risikoreduktion), als auch dessen Nachteile (falsche Alarme, Überdiagnosen) werden auf eine und dieselbe fiktive Grundgesamtheit bezogen und dadurch unmittelbar vergleichbar (siehe Abbildung 2). Vielleicht darf an dieser Stelle hinzugefügt werden, dass die Autoren dieser Darstellungen mit den Ergebnissen von Studien wie zum Beispiel den oben berichteten gut vertraut waren.

Die in Abbildung 2 gegebene Übersicht fasst der Autor in Worte ([11], S. 23), die in eine für die Allgemeinheit verfasste Informationsbroschüre wie folgt übernommen worden sind ([20], S. 10).

Folgende Zahlen, die auf Erfahrungen aus anderen Ländern und auf wissenschaftlichen Untersuchungen beruhen, sollen Ihnen eine konkrete Vorstellung davon geben, wie Vor- und Nachteile über das gesamte Programm in etwa statistisch verteilt sind:

- Von 200 Frauen, die 20 Jahre lang jedes 2. Jahr am Mammographie-Screening-Programm teilnehmen, erhalten 140 in 20 Jahren keinen verdächtigen Befund. 60 Frauen bekommen einen Befund, dem nachgegangen werden sollte.
- Von diesen 60 Frauen erhalten 40 bei der ergänzenden Untersuchung Entwarnung, 20 Frauen wird eine Gewebeentnahme empfohlen.
- Von diesen 20 Frauen stellt sich bei 10 Frauen der Verdacht als unbegründet heraus. 10 Frauen erhalten die Diagnose Brustkrebs im Screening. Von den übrigen 190 Frauen erhalten drei Frauen in den 20 Jahren zwischen zwei Screeningrunden ebenfalls die Diagnose Brustkrebs.
- Von diesen insgesamt 13 Frauen mit der Diagnose Brustkrebs sterben drei Frauen an Brustkrebs, 10 Frauen sterben nicht an Brustkrebs.
- Von diesen 10 Frauen hätte eine Frau ohne Mammographie zu Lebzeiten nichts von ihrem Brustkrebs erfahren, acht Frauen wären auch ohne Teilnahme am Mammographie-Screening-Programm erfolgreich behandelt worden – ein Teil davon jedoch mit einer belastenderen Therapie. 1 von 200 Frauen wird dank ihrer regelmäßigen Teilnahme vor dem Tod durch Brustkrebs bewahrt.

Diese Zusammenstellung erfüllt alle Kriterien des Katalogs, den der Fachbereich Patienteninformatik des Deutschen Netzwerks für evidenzbasierte Medizin (DNEbM) erstellt hat, um Ärzte bei der Pflichtberatung

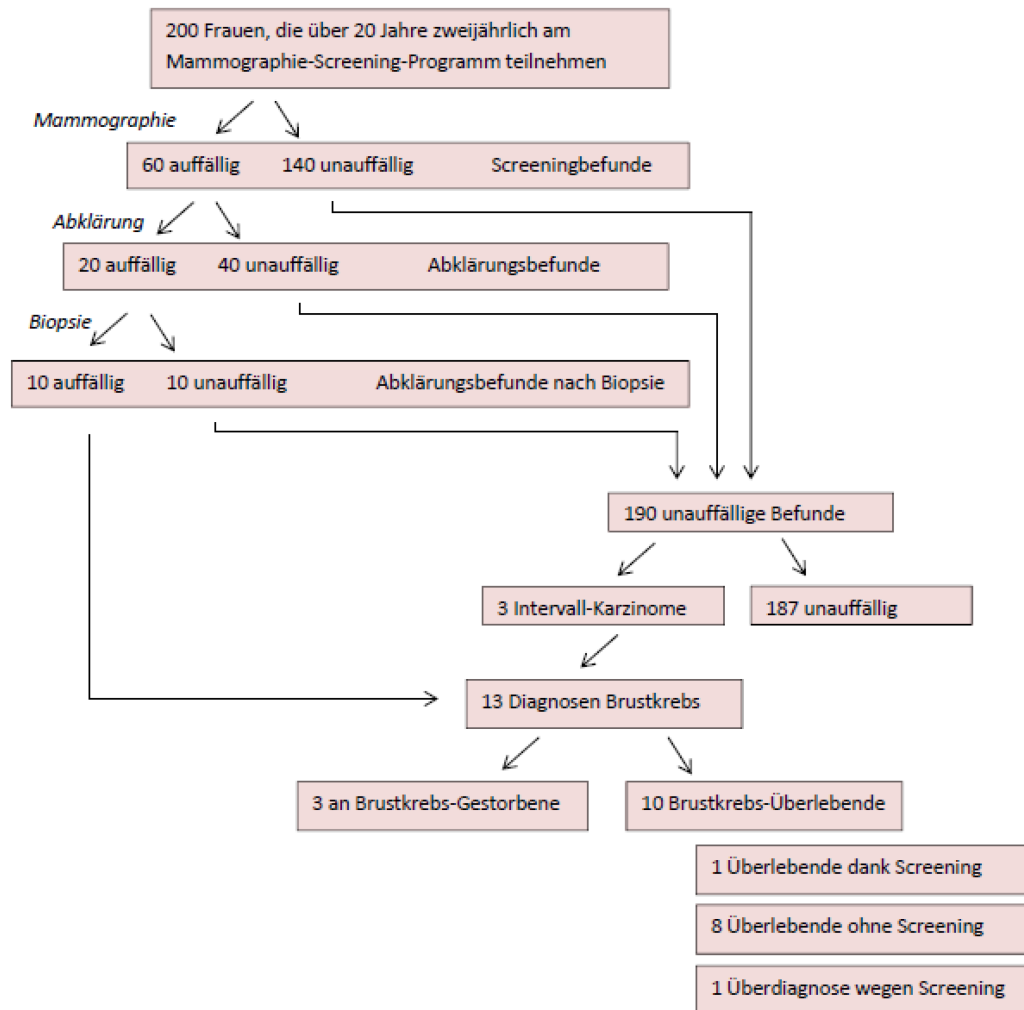


Abbildung 2: Nutzen der Mammographie (adaptiert aus [11])

zur Krebsfrüherkennung zu unterstützen [21] (siehe auch [22]) und sie entspricht den Forderungen nach einen besseren Risikokommunikation im Kontext von Screeningprogrammen [23]. Sie ist klar und transparent, und kommuniziert die wichtigsten Zahlen auf eine Art und Weise, die es erlaubt, Nutzen und Nachteile direkt zu vergleichen. So ist jede Frau selbstständig oder auch im Gespräch mit ihrem Arzt in der Lage zu entscheiden, ob sie an dieser Lotterie – genannt Mammographie-Screening – teilnehmen möchte (für weitere lobenswerte Darstellungen, siehe [24], [25], [26]). Wir möchten Ärzte, Fachgesellschaften, Patientenorganisationen, Krankenkassen und Behörden dazu ermuntern, dieses Beispiel aufzugreifen und weitere derart transparente Übersichten zu Diagnose und Therapie diverser Erkrankungen zu erstellen. Das Harding Center für Risikokompetenz (Risk Literacy) am Max-Planck Institut für Bildungsforschung in Berlin nennt diese Übersichten „Faktenboxen“ (siehe auch [27], [28]) und hat auch bereits schon eine Reihe davon vorgelegt [29]. Risiken sind unvermeidbar. Es gab sie schon immer und wird sie auch weiterhin geben – aber schlechte Risikokommunikation und Missverständnisse... das muss nun wirklich nicht sein.

Anmerkungen

Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Literatur

1. Ahmed H, Naik G, Willoughby H, Edwards AG. Communicating risk. *BMJ*. 2012;344:e3996. DOI: 10.1136/bmj.e3996
2. Gigerenzer G, Muir Gray JA, editors. Better doctors, better patients, better decisions: Envisioning health care 2020. Cambridge: MIT; 2011.
3. Malek D, Käab-Sanyal V; Kooperationsgemeinschaft Mammographie. Evaluationsbericht 2011 – Zusammenfassung der Ergebnisse des Mammographie-Screening-Programms in Deutschland. Berlin: Kooperationsgemeinschaft Mammographie; 2014. Available from: http://www.mammo-programm.de/download/MAMMO_EvalBericht_20141201_web%282%29.pdf [cited 2015 Mar 8]

4. Eddy DM. Probabilistic reasoning in clinical medicine: Problems and opportunities. In: Kahneman D, Slovic P, Tversky A, editors. *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press; 1982. pp. 249-67. DOI: 10.1017/CBO9780511809477.019
5. Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Acad Med*. 1998 May;73(5):538-40. DOI: 10.1097/00001888-199805000-00024
6. Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G. Medicine. Communicating statistical information. *Science*. 2000 Dec;290(5500):2261-2. DOI: 10.1126/science.290.5500.2261
7. Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychol Rev*. 1995;102:684-704. DOI: 10.1037/0033-295X.102.4.684
8. Ghosh AK, Ghosh K. Translating evidence-based information into effective risk communication: current challenges and opportunities. *J Lab Clin Med*. 2005 Apr;145(4):171-80. DOI: 10.1016/j.lab.2005.02.006
9. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients to make sense of health statistics. *Psychol Sci Public Interest*. 2007;8:53-96. DOI: 10.1111/j.1539-6053.2008.00033.x
10. Kurzenhäuser S, Hoffrage U. Teaching Bayesian reasoning: an evaluation of a classroom tutorial for medical students. *Med Teach*. 2002 Sep;24(5):516-21. DOI: 10.1080/0142159021000012540
11. Weymayr C; Kooperationsgemeinschaft Mammographie. Kennzahlen Mammographie-Screening – Dokumentation 2010 Version 1.2. Berlin: Kooperationsgemeinschaft Mammographie; 2010. Available from: <http://www.komen.de/html/img/pool/1kennzahlenmammographie-screeningdokumentationv1.2.pdf> [cited 2015 Mar 8]
12. Kurzenhäuser S. Welche Informationen vermitteln deutsche Gesundheitsbroschüren über die Screening-Mammographie [What kind of information do German health information pamphlets provide on mammography screening?]. *Z Arztl Fortbild Qualitatssich*. 2003 Feb;97(1):53-7.
13. Kurzenhäuser S, Hoffrage U. Designing risk communication in health. In: Todd PM, Gigerenzer G; ABC Research Group, editors. *Ecological rationality – intelligence in the world*. New York: Oxford University Press; 2012. p. 428-53.
14. Woloshin S, Schwartz LM. Numbers needed to decide. *J Natl Cancer Inst*. 2009 Sep;101(17):1163-5. DOI: 10.1093/jnci/djp263
15. Matter-Walstra K, Hoffrage U. Individuelle Entscheidungsfindung am Beispiel der Brustkrebs-Früherkennung—Erfahrungen aus Fokusgruppen in der Schweiz. [Individual decision making concerning breast cancer screening—Observations with focus groups in Switzerland]. *Schweizer Zeitschr Managed Care Care Management*. 2001;3:26-9.
16. Wegwarth O, Gigerenzer G. Risikokommunikation – Risiken und Unsicherheiten richtig verstehen lernen. *Dtsch Arztebl*. 2011;108(9):A-448-51.
17. Sökeland J, Weiss HE. Häufigkeit des Prostatakarzinoms. 2011 [last update 2014 Jul 28]. Available from: http://www.prostata.de/pca_haeufigkeit.html [cited 2015 Mar 8]
18. Wegwarth O, Gaissmaier W, Gigerenzer G. Deceiving numbers: survival rates and their impact on doctors' risk communication. *Med Decis Making*. 2011 May-Jun;31(3):386-94. DOI: 10.1177/0272989X10391469
19. Welch HG, Frankel BA. Likelihood that a woman with screen-detected breast cancer has had her "life saved" by that screening. *Arch Intern Med*. 2011 Dec;171(22):2043-6. DOI: 10.1001/archinternmed.2011.476
20. Gemeinsamer Bundesausschuss. Informationen zum Mammographiescreening. Available from: http://www.mammo-programm.de/download/merkblatt_deutsch_web.pdf [cited 2015 Mar 8]
21. Koch K, Mühlhauser I. Kriterien zur Erstellung von Patienteninformationen zu Krebsfrüherkennungsuntersuchungen: Stellungnahme des Fachbereichs Patienteninformation des Deutschen Netzwerkes für Evidenzbasierte Medizin (DNEbM). 2008. Available from: <http://www.ebm-netzwerk.de/pdf/stellungnahmen/dnebm-080630.pdf> [cited 2015 Mar 8]
22. Griebenow B. Beratung zur Krebsfrüherkennung: Vor- und Nachteile darstellen. *Dtsch Arztebl*. 2008;05(33):A-1724.
23. Jørgensen KJ, Brodersen J, Hartling OJ, Nielsen M, Gøtzsche PC. Informed choice requires information about both benefits and harms. *J Med Ethics*. 2009 Apr;35(4):268-9. DOI: 10.1136/jme.2008.027961
24. Leitlinienprogramm Onkologie der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V., der Deutschen Krebsgesellschaft e. V. und der Deutschen Krebshilfe e. V. Früherkennung von Brustkrebs – Eine Entscheidungshilfe für Frauen. 2010. Available from: http://leitlinienprogramm-onkologie.de/uploads/tx_sbdownloader/Patientenleitlinie_Brustkrebs_Fruherkennung.pdf [cited 2014 Nov 25]
25. Medizinische Universität Graz, EBM Review Center. Mammographiebasierte Brustkrebsfrüherkennung – Recherche und Aufbereitung von Kennzahlen für eine informierte Entscheidung. Available from: http://www.frueh-erkennen.at/tl_files/bkfp/Dokumente/Fuer%20Radiologen/Mammographie%20basierte%20Screening-Kennzahlen%20informierte%20Entscheidung.pdf [cited 2015 Mar 8]
26. Gøtzsche PC, Hartling OJ, Nielsen M, Brodersen J. Screening für Brustkrebs mit Mammographie. Copenhagen: Nordisches Cochrane Zentrum; 2012. Available from: <http://www.cochrane.dk/screening/mammografi-de.pdf> [cited 2015 Mar 8]
27. Schwartz LM, Woloshin S, Welch HG. The drug facts box: providing consumers with simple tabular data on drug benefit and harm. *Med Decis Making*. 2007 Sep-Oct;27(5):655-62. DOI: 10.1177/0272989X07306786
28. Schwartz LM, Woloshin S, Welch HG. Using a drug facts box to communicate drug benefits and harms: two randomized trials. *Ann Intern Med*. 2009 Apr;150(8):516-27. DOI: 10.7326/0003-4819-150-8-200904210-00106
29. Harding Zentrum für Risikokompetenz. Faktenboxen. Available from: <https://www.harding-center.mpg.de/de/gesundheitsinformationen/faktenboxen> [cited 2015 Mar 8]
30. Hoffrage U, Kurzenhäuser S, Gigerenzer G. Wie kann man die Bedeutung medizinischer Testbefunde besser verstehen und kommunizieren [How can one improve the understanding and communication of the importance of medical test results?]. *Z Arztl Fortbild Qualitatssich*. 2000 Oct;94(9):713-9.

Korrespondenzadresse:

Prof. Dr. Ulrich Hoffrage
Faculty of Business and Economics, University of
Lausanne, 1015 Lausanne, Switzerland, Phone: +41 21
692 3490, Fax: +41 21 692 3305
Ulrich.hoffrage@unil.ch

Bitte zitieren als

Hoffrage U, Koller M. Chances and risks in medical risk communication. *GMS Ger Med Sci*. 2015;13:Doc07.
DOI: 10.3205/000211, URN: urn:nbn:de:0183-0002117

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/gms/2015-13/000211.shtml>

Eingereicht: 02.12.2014

Überarbeitet: 04.05.2015

Veröffentlicht: 06.07.2015

Copyright

©2015 Hoffrage et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.