

---

Ute Goerling · Anja Mehnert  
Editors

# Psycho-Oncology

Second Edition

*Editors*

Ute Goerling  
Charité Comprehensive Cancer Center  
Humboldt University of Berlin  
Berlin  
Germany

Anja Mehnert  
Department of Medical Psychology  
and Medical Sociology  
University of Leipzig  
Leipzig  
Germany

ISSN 0080-0015

ISSN 2197-6767 (electronic)

Recent Results in Cancer Research

ISBN 978-3-319-64309-0

ISBN 978-3-319-64310-6 (eBook)

<https://doi.org/10.1007/978-3-319-64310-6>

Library of Congress Control Number: 2013953265

© Springer International Publishing AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland



---

# The Barrier to Informed Choice in Cancer Screening: Statistical Illiteracy in Physicians and Patients

Odette Wegwarth and Gerd Gigerenzer

---

## 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 nine European countries, 92% overestimated the reduction of breast cancer mortality by mammography by a factor of 10–200, or did not know. For a similar sample of about 5000 men with respect to PSA screening, this number was 89%. Of more than 300 US citizens who regularly attended one or more cancer screening test, 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, with estimates varying between 1 and 90%. In a national sample of 412 US primary care physicians, 47% mistakenly believed that if more cancers are detected by a screening test, this proves that the test saves lives, and 76% wrongly thought that if screen-detected cancers have better 5-year survival rates than cancers detected by symptoms, this would prove that the screening test saves lives. And of 20 German gynecologists, not a single one provided a woman with all information on the benefits and harms of cancer screening required in order to make an informed choice. Why is risk literacy so scarce in health care? One frequently discussed explanation assumes that people suffer from cognitive deficits that make them predictably irrational and basically hopeless at dealing with risks, so that they need to be “nudged” into healthy behavior. Yet research has

---

O. Wegwarth (✉) · G. Gigerenzer  
Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany  
e-mail: wegwarth@mpib-berlin.mpg.de

© Springer International Publishing AG 2018  
U. Goerling and A. Mehnert (eds.), *Psycho-Oncology, Recent Results*  
in Cancer Research 210, [https://doi.org/10.1007/978-3-319-64310-6\\_13](https://doi.org/10.1007/978-3-319-64310-6_13)

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 uses relative risks and other misleading statistics, motivated by conflicts of interest and defensive medicine that do not promote informed physicians and patients. 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—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**

Informed decision-making · Cancer screening · Medical risk illiteracy · Absolute risk · Relative risk · 5-year survival · Medical risk communication

---

## 1 Introduction

Patients appear to be the problem in modern high-tech health care: uninformed, anxious, noncompliant, and with unhealthy lifestyles. They demand drugs advertised by celebrities on TV, insist on unnecessary but expensive tests and treatments, and may eventually turn into plaintiffs. In light of skyrocketing health costs in Western countries, patients' lack of health literacy and the resulting costs and harms have received much attention. Consider a few cases. Almost 10 million U.S. women have had unnecessary Pap smears to screen for cervical cancer—unnecessary because each of them had had a complete hysterectomy and thus no longer had a cervix (Sirovich and Welch 2004). Unnecessary Pap tests may cause no harm but they waste millions that could be put to better use in health care. Every year, one million U.S. children have unnecessary computed tomography (CT) scans (Brenner and Hall 2007). An unnecessary CT scan is more than a waste of money; an estimated 29,000 cancers result from the approximately 70 million CT scans performed annually in the U.S. (Berrington de González et al. 2009). And when a random sample of 500 Americans was asked whether they would rather receive one thousand dollars in cash or a free full-body CT, three out of four wanted the CT (Schwartz et al. 2004). Why do people not protect their children or themselves from unnecessary doses of radiation? They probably would if they knew the risks involved. Uninformed patients are by no means restricted to the U.S. A representative study of over 10,000 citizens in nine European countries revealed that 89% of

men and 92% of women, respectively, overestimated the benefit of PSA and mammography screening 10-fold, 100-fold and more, or did not know it (Gigerenzer et al. 2009). With more and better access to health information than ever, why are people so largely uninformed?

The answers proposed include that many patients are not intelligent enough or do not want to deal with numbers, even though most 12-year-olds in the U.S. already know baseball statistics, and their British peers can recite the relevant numbers of the Football Association (FA) Cup results. Scores of health psychologists and behavioral economists add to the list of suspected cognitive deficits by emphasizing patients' cognitive biases, weakness of will, and wishful thinking (Gigerenzer and Gray 2011). In this view, the problems facing health care are people who engage in self-harming behavior, focus on short-term gratification rather than long-term harms, suffer from the inability to predict their emotional states after a treatment, or simply do not want to think and prefer to trust their doctor (Wegwarth and Gigerenzer 2013). Consequently, the recommended remedies are some form of paternalism that "nudges" the inept patient in the right direction (Thaler and Sunstein 2008). Yet the most decisive reason for the lack of health literacy in patients is far more likely the widespread amount of misinformation, whose sources are risk illiterate physicians, intransparent patients' brochures, and the media.

---

## 2 How to Communicate the Benefit and Harms of Cancer Screening?

Imagine a 53-year-old woman who considers attending mammography screening for breast cancer. To make an informed decision about whether to attend, she needs to learn about the benefits and harms of that cancer screening. What is the current evidence? In 1996, results of four randomized trials on mammography screening including approximately 280,000 women (Nyström et al. 1996) showed that of 1000 women attending screening over 10 years, three women died of breast cancer, whereas of 1000 women not attending screening over 10 years, four women died of breast cancer. Further analysis showed similar effects: Here, breast cancer mortality was 4 out of 1000 women who attended mammography screening over a course of 10 years, compared to 5 out of 1000 who did not (Nyström et al. 2002). Thus, in both analyses the absolute reduction of breast cancer death due to mammography was 1 woman in 1000. Subsequent Cochrane reviews of these and further randomized trials enrolling approximately 500,000 women found the absolute risk reduction to be even smaller: Now it was estimated that mammography screening would save only 1 woman in 2000 (Gøtzsche and Nielsen 2006, 2011) from breast cancer death. In addition, authors quantified mammography's harms: overdiagnosis and overtreatment. Overdiagnosis refers to the detection of pseudodisease—screening-detected abnormalities that meet the pathologic definition of cancer but will never progress to cause symptoms or cancer death in the patient's lifetime. The consequence of

overdiagnosis is overtreatment—surgery, chemotherapy, or radiation that provides the patient with no survival benefit but only side effects. For mammography it is estimated that for every 2000 women invited for screening throughout 10 years, 10 women who would not have been diagnosed with breast cancer if they had not been screened will be treated unnecessarily. Furthermore, more than 200 women out of these 2000 will experience important psychological distress, including anxiety and uncertainty for years because of false-positive findings.

Is the woman in question likely to learn about that evidence from her physician? To learn more about gynecologists' counseling on mammography, we conducted a study (Wegwarth and Gigerenzer 2011) in 2008, nearly 2 years after the first comprehensive Cochrane review about the benefits and harms of mammography was published (Gøtzsche and Nielsen 2006, 2011). One of us called gynecologists who were practicing in different cities across Germany and told them the following story: Our 55-year-old mother with no history of breast cancer in her family and without any symptoms had received an invitation to attend a mammography screening but doubted its effectiveness; we, in contrast, believed that it might be advisable to attend and would like to learn in more detail about its benefits and harms. Of the 20 gynecologists who were willing to talk to us, 17 strongly recommended mammography, emphasizing that it is a safe and scientifically well-grounded intervention. Only seven of these were able to provide numbers for the requested benefit of a reduced risk of breast cancer death, which ranged from 20 to 50%. Communication of the harms was even more discouraging: None of the gynecologists mentioned the risk of being overdiagnosed or overtreated as a consequence of mammography screening. Instead, the majority described the potential harms as "negligible" and "harmless." Only 3 out of the 20 gynecologists provided numbers for specific harms, out of which two numbers were wrong.

The results of these studies documented two issues: (1) People who consult their physicians on the benefits and harms of cancer screening are unlikely to receive correct numbers, if any, on the benefits and harms but instead verbal and subjective qualifiers, and (2) they are likely to be misled by *mismatched framing* (Gigerenzer et al. 2007). *Mismatched framing* refers to the act of reporting the benefits and harms of a medical intervention in different "currencies": usually the benefits in relative risks (=large numbers) and the harms in absolute risks (=small numbers). The same risk reduction (for benefits) or risk increase (for harms) can be expressed as either a relative risk (RR), absolute risk (AR), or the number of people needed to be treated (screened) to prevent one death from cancer (NNT, which is  $1/\text{absolute risk reduction}$ ). For instance, taking the review on mammography screening (Gøtzsche and Nielsen 2006), where a breast cancer mortality reduction from 5 to 4 women in 2000 was observed, one can report these results as

**RRR** If you have regular mammography screening, it will reduce your chances of dying from this cancer by around 20% over the next 10 years.

**ARR** If you have regular mammography screening, it will reduce your chances of dying from this cancer from around 5 per 2000 to around 4 per 2000 over the next 10 years.



**NNT** To save 1 woman from dying from breast cancer over the next 10 years, around 2000 women have to have regular mammography screening.

Whereas absolute risks and NNT are typically small numbers, the corresponding relative risk tends to appear large. As a consequence, the format of relative risk leads not only laypersons but also doctors to overestimate the benefits of medical interventions. In our study about counseling on mammography screening, all numerical information we received from the gynecologists about the benefit were relative risk reductions, whereas the harms were quantified as absolute risk increase. For instance, the estimates we received for the benefit (reduction of breast cancer deaths) ranged from 20 to 50%. The 50% does not correspond to any findings of evidence-based studies on the effectiveness of mammography screening, but the 20% corresponds to results of earlier reviews (Nyström et al. 2002). To arrive at the 20%, all other information (e.g., how many women were in each of the studied groups = reference classes) is ignored and only the reduction from five breast cancer deaths (=100%) to four breast cancer deaths (=80% from 5) is considered. What this relative risk statement suggests to most readers is that of all people who are screened, 20% fewer die of breast cancer. Yet that is not what the 20% means. In fact, a relative risk of 20% can be compatible with a wide range of changes in the absolute risk reduction of death, such as a reduction from 50 to 40, from 1000 to 800, and from 0.0005 to 0.0004. Without specifying the underlying absolute risks, i.e., the absolute numbers of breast cancer deaths in the screening group and the non-screening group, as well as the sample size of each of the groups, the information is incomplete (Forrow et al. 1992). Effects presented in relative terms thus communicate very little about the true and absolute size of the effect of the medical mean.

---

### **3 Why Is the Misleading Relative Risk Information so Commonly Used?**

As mentioned earlier, relative risk information typically yields large numbers and absolute risk information small numbers. This means that relative risk information appears much more impressive to physicians (Fahey et al. 1995; Naylor et al. 1992), policy makers (Hux and Naylor 1995), and patients (Malenka et al. 1993; Schwartz et al. 1997). For instance, in a study in a Swiss hospital, 15 gynecologists were asked what the widely known 25% risk reduction through mammography actually means (Schüssler 2005). This number corresponds to the first review released in 1996 (Nyström et al. 1996) on the effects of mammography attendance, where the risk of dying from breast cancer was reduced from 4 to 3 (=25%) women in 1000. Asked how many fewer women die of breast cancer given the relative risk reduction of 25%, one physician thought that 25% meant 2.5 out of 1000, another 25 out of 1000; the total range of the answers was between 1 and 750 in 1000 women. At the beginning of a CME course in risk communication, another group of 150

gynecologists was also asked what the 25% risk figure meant (Gigerenzer et al. 2007). Using an interactive voting system, the physicians could choose between four alternatives

Mammography screening reduces mortality from breast cancer by about 25%. Assume that 1000 women aged 40 and over participate in mammography screening. How many fewer women are likely to die of breast cancer?

1 [66%]

25 [16%]

100 [3%]

250 [15%]

The numbers in brackets show the percentage of gynecologists who gave the respective answer. Two-thirds understood that the best answer was 1 in 1000. Yet 16% believed that the figure meant 25 in 1000, and 15% responded that 250 fewer women in 1000 would die of breast cancer.

Where does this confusion come from? Next to the fact that more than 90% of all medical research is financed by the pharmaceutical industry—which has an obvious interest in making results look good—medical journals, even high-ranking ones, also play a role in spreading intransparent statistics. Studies on the coverage of medical findings in high-ranking medical journals revealed that nontransparent health statistics such as relative risk reduction are the rule rather than the exception. In their analysis of 359 articles that reported randomized trials in the years 1989, 1992, 1995, and 1998, published in *Annals of Internal Medicine*, *British Medical Journal (BMJ)*, *Journal of the American Medical Association (JAMA)*, *The Lancet*, and *The New England Journal of Medicine* Nuovo et al. (2002) found that only 25 articles reported absolute risk reduction and 14 of these 25 also included *the number needed to treat (NNT)*, which is simply the inverse of the absolute risk reduction. That is, only about 7% of the articles reported the results in a transparent way. The same journals, along with the *Journal of the National Cancer Institute*, were analyzed again in 2003/2004 (Schwartz et al. 2006). Sixty-eight percent of 222 articles failed to report the absolute risks for the first ratio measure (such as relative risks) in the abstract, and about half of these did not report the underlying absolute risks anywhere at all in the article. An analysis of *BMJ*, *JAMA*, and *The Lancet* from 2004 to 2006 found that in about half of the articles, absolute risks or other transparent frequency data were not reported (Sedrakyan and Shih 2007). The study further revealed that 1 out of 3 studies used mismatched framing when reporting their findings. In most cases, relative risks (=large numbers) were reported for benefits, and absolute risks (=small numbers) for harms. In 2010, we sought to find out whether the situation had since changed and investigated all free available research articles reporting drug interventions published in *BMJ* in 2009 (Gigerenzer et al. 2010). Of the 37 articles identified, 16 failed to report the underlying absolute numbers for the reported relative risk measures in the abstract. Among these, 14 reported the absolute risks elsewhere in the article, but 2 did not report them at all. Moreover, absolute risks or number needed to treat (NNT) were more often

reported for harms ( $10/16 = 63\%$ ) than for benefits ( $14/27 = 52\%$ ). These analyses indicate that one reason why physicians, patients, and journalists talk about relative risk reductions is because most of the original studies regularly provide information in this form.

Leaflets—developed by the pharmaceutical industry to inform doctors and patients of medical products, tests, and treatments—are even worse. Comparing the summaries in 175 leaflets with the original studies (Kaiser et al. 2004), researchers from the German Institute for Quality and Efficiency in Health Care (IQWiG) found that in only 8% of the cases could the summaries be verified. In the remaining 92%, key results of the original study were systematically distorted or important details omitted. For instance, one pamphlet from Bayer stated that their potency drug Levitra (Vardenafil) works up to 5 h—without mentioning that this statistic was based on studies with numbed hares. Moreover, the cited sources were often either not provided or impossible to find. In general, leaflets exaggerated baseline risks and risk reduction, enlarged the period in which medication could safely be taken, or did not reveal severe side effects of medication pointed out in the original publications.

---

#### 4 What Is the Lesson to Be Learned from This?

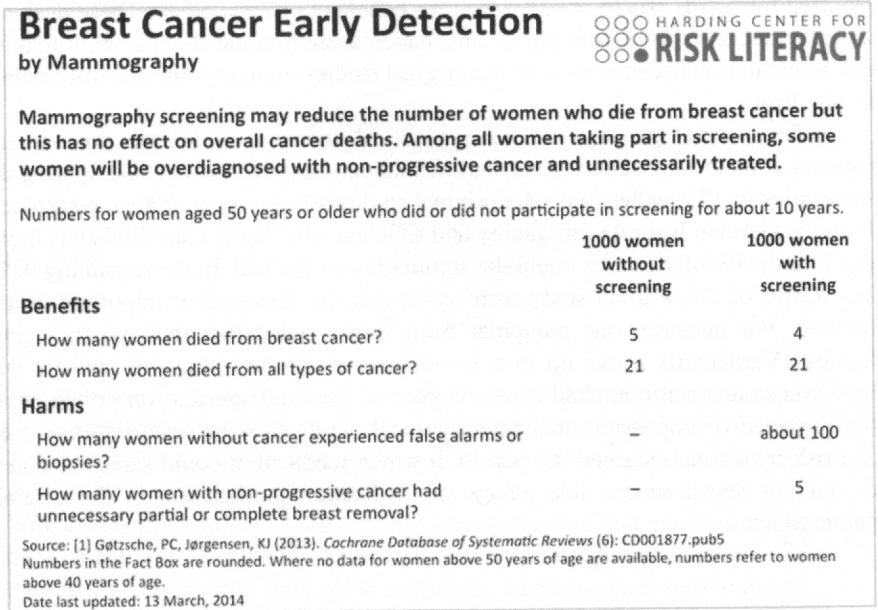
First, always be aware that not only treatments and drugs but also screening tests have benefits *and* harms. Second, when judging or communicating screening's benefits and harms, do not rely on percentages or ratio measures. Instead, find the absolute numbers of people involved in the intervention group (here, screening group) and control group (here, non-screening group) and the absolute numbers of the event (e.g., number of cancer deaths) in both groups. Third, to make the benefits and harms comparable to each other, adjust the numbers of events to the same and smallest possible denominator (e.g., 1000 people). The following fact box on mammography provides a good example of transparent risk communication of benefits and harms (Fig. 1).

---

#### 5 Does a Positive Test Result of Cancer Screening Mean Having Cancer for Certain?

Doctors' understanding of a positive and a negative test result is essential for a patient who has taken a test. Not knowing and thereby miscommunicating the meaning of a positive result can lead to overdiagnosis, overtreatment, unnecessary fear, or sometimes even to suicide.

*Mammography:* Consider a woman who has just received a positive mammogram and who asks her doctor whether she has breast cancer for certain, and if not, what the chances are. One would assume that every gynecologist knows the answer.



**Fig. 1** Fact box on effectiveness of mammography screening

But does the assumption hold true? At the beginning of one continuing education session, 160 gynecologists (Gigerenzer et al. 2007) were provided with the relevant health statistics needed for answering this question in the form of conditional probabilities, which is the form in which medical studies tend to report these health statistics

Assume that you conduct breast cancer screening using mammography in a certain region. You know the following information about the women in this region:

- (1) The probability that a woman has breast cancer is 1% (prevalence).
- (2) If a woman has breast cancer, the probability that she tests positive is 90% (sensitivity).
- (3) If a woman does not have breast cancer, the probability that she nevertheless tests positive is 9% (false-positive rate).

A woman tests positive. She wants to know from you whether this means that she has breast cancer for sure, or what the chances are. What is the best answer?

- (A) The probability that she has breast cancer is about 81%.
- (B) Out of 100 women with a positive mammogram, about 90 have breast cancer (90%).
- (C) Out of 100 women with a positive mammogram, about 10 have breast cancer (10%).
- (D) The probability that she has breast cancer is about 1%.

Gynecologists could either derive the answer from the health statistics provided or simply recall what they should have known anyhow. In either case, the best answer is (C). That is, only about 10 out of every 100 women who test positive in



screening actually has breast cancer. The other 90 women are falsely alarmed, that is, have a false-positive test result because they do not have breast cancer. Only 21% of the gynecologists found the best answer; the majority (60%) disconcertingly chose the options of "90%" or "81%," thus grossly overestimating the probability of cancer. Another troubling result was the high variability in physicians' estimates, ranging between a 1% and 90% chance of cancer.

*Fecal occult blood test (FOBT) screening:* Hoffrage and Gigerenzer (1998) tested 48 physicians with an average professional experience of 14 years, including radiologists, internists, surgeons, urologists, and gynecologists. The sample had physicians from teaching hospitals slightly overrepresented and included heads of medical departments. They were given four problems, one of which concerned screening for colorectal cancer with the fecal occult blood test. Half of the physicians were given the relevant information in conditional probabilities (a sensitivity of 50%, a false-positive rate of 3%, and a prevalence of 0.3%). This group of physicians was then asked to estimate the probability of colorectal cancer given a positive test result. Their estimates ranged from a 1 to 99% chance of cancer after a positive test. Their modal answer was 50% (the sensitivity); four physicians deducted the false-positive rate from the sensitivity (arriving at 47%). When interviewed about how they arrived at their answers, several physicians claimed to be innumerate and hid this from patients by avoiding any mention of numbers.

Already back in 1978, a study (Casscells et al. 1978) documented that the majority of physicians struggled with making correct inferences from positive test results: Only 18% of the physicians and medical staff who participated could correctly infer the likelihood of having a disease given a positive test result (positive predictive value/PPV) from the given information. Somewhat later Eddy (1982) reported that 95 out of 100 physicians overestimated the probability of cancer after a positive screening mammogram by an order of magnitude. Similarly, Bramwell et al. (2006) found that only 1 out of 21 obstetricians was able to estimate the probability of an unborn child actually having Down Syndrome given a positive test, with those giving incorrect responses being fairly confident in their estimates. And in an Australian study, 13 of 50 physicians claimed they could describe the positive predictive value of a test, yet when directly interviewed, only 1 could do so (Young et al. 2002). Similar effects were reported for members of the U.S. National Academy of Neuropsychology (Labarge et al. 2003). Ghosh and Ghosh (2005) reviewed further studies that showed that few physicians were able to estimate the positive predictive value from the relevant health statistics.

---

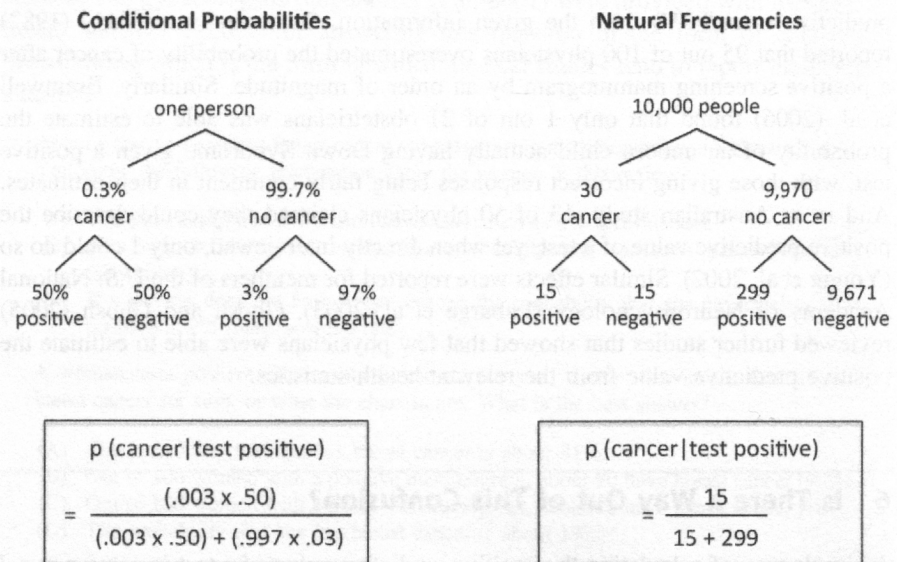
## 6 Is There a Way Out of This Confusion?

A simple way of calculating the positive predictive value of a test is using *natural frequencies* (Gigerenzer and Hoffrage 1995). The principle of this approach rests on the assumption that our brains are shaped by evolution to the use of naturally gathered frequencies rather than probabilities, which were unknown before the late

eighteenth century. Using that approach of natural frequencies entails “translating” all of the probabilistic information into frequencies. To illustrate, consider the example of the screening for colorectal cancer with the fecal occult blood test (FOBT) again with a sensitivity of 50%, a false-positive rate of 3%, and a prevalence of 0.3%.

- *Prevalence 0.3%:* Out of 10,000 people about 30 people actually have colorectal cancer. (0.3% of 10,000)
- *Sensitivity 50%:* Of these 30 people with colorectal cancer 15 will receive a true positive test result. (50% of 30) (The other 15 will receive a false negative result).
- *False-positive rate 3%:* Of the 9,970 people without colorectal cancer (10,000 minus 30 with colorectal cancer), 299 will receive a false-positive test result. (3% of 9970)

Altogether 314 people will receive a positive test result by the test, but for only 15 people is it correct. Thus, the likelihood of a person having colorectal cancer if their FOBT test is positive is about 5%. Given that physicians’ estimates ranged from a 1 to 99% chance of colorectal cancer after a positive test with a modal answer of 50%, one can easily imagine how many patients will be unnecessarily frightened as a byproduct of their physicians’ statistical illiteracy. Figure 2 illustrates the calculation within a natural frequency tree.



**Fig. 2** The probability of colorectal cancer given a positive fecal blood test result. The *left* side illustrates the calculation with conditional probabilities, while the *right* side provides a more transparent calculation with natural frequencies

## 7 Does an Increase in 5-Year Survival Rates Mean that Lives Are Saved?

While running for president, Rudi Giuliani, former New York City mayor, said in a 2007 campaign advertisement:

I had prostate cancer, 5, 6 years ago. My chance of surviving prostate cancer—and thank God, I was cured of it—in the United States? Eighty-two percent. My chance of surviving prostate cancer in England? Only 44 percent under socialized medicine (Dobbs 2007, October 30).

This difference in 5-year survival between the U.S. and the U.K. appears large. But is it really that different? It is not, although most people will not realize that they were misled by Giuliani. Giuliani presented higher 5-year survival rates as suggestive evidence for lower mortality due to screening, when in fact differences in survival rates are *uncorrelated* with differences in mortality rates (Welch et al. 2000). In reality, mortality from prostate cancer is about the same in the U.S. and the U.K., even though most American men take the PSA (Prostate-specific Antigen) test and most British men do not. There are two reasons why higher survival rates tell us nothing about lower mortality in the context of screening. First, screening results in early detection and thus inflates 5-year survival rates by simply setting the point of diagnosis earlier, without necessarily extending life (*lead time bias*). As a consequence, people may just live earlier (and longer) with the diagnosis than do people whose cancer is detected by symptoms but die no later than do people diagnosed by symptoms. As for the 5-year survival rate the clock starts ticking at the moment of the diagnosis, and thus people in the screening group are more likely to be still alive 5 years after the earlier diagnosis. Yet that does not mean that they have gained a single extra month of life compared to people without screening. Second, screening inflates survival rates by including people with non-progressive cancers that by definition do not lead to death (*overdiagnosis bias*; Gigerenzer et al. 2007). As a consequence the ratio between the number of people diagnosed with cancer (including the non-progressive types) and the number of diagnosed people still alive after 5 years automatically looks more favorable. Giuliani is not the only one to have misled the public with survival rates; other guilty parties include prestigious U.S. cancer centers such as MD Anderson (Gigerenzer et al. 2007) and high-profile charities such as The Susan G. Komen Association (Woloshin and Schwartz 2012).

---

## 8 What Do Physicians Know About the 5-Year Survival Statistic in the Context of Screening?

One might think that physicians would provide people with the right numbers precisely in order to avoid such misunderstandings and facilitate informed choice. Yet studies document that this is unlikely to happen. Few doctors themselves are

aware that in screening, survival rates tell us nothing about mortality; nor do they know what lead time bias and overdiagnosis bias are (Wegwarth et al. 2011, 2012). More specifically, in a national sample of 412 US primary care physicians, 47% wrongly thought that if more cancers are detected by a screening test, this proves that the test saves lives, and 76% mistakenly believed that if screen-detected cancers have better 5-year survival rates than cancers detected by symptoms, this too proves that a test saves lives (Wegwarth et al. 2012). When provided with data on what appeared to be two screening tests, primary care physicians were more enthusiastic about the test supported by an increase in 5-year survival (increase of 31 percentage points) than about the test supported by reduced cancer mortality (reduction of 0.4 men in 1000): 69 versus 23%, respectively of the very same physicians said they would definitely recommend the test to their patients. In fact, all data came from medical evidence on the same cancer screening test—prostate cancer screening. These results demonstrate not only that physicians do not correctly understand cancer screening statistics but—even worse—that 46% of the physicians in our sample would have given their patients conflicting advice about a single cancer screening procedure, depending on what statistics they were confronted with.

If physicians do not understand medical statistics, clearly they cannot support informed decision-making in their patients. And if their physicians are of little help, do patients have a chance of making an informed choice after reading patient pamphlets or media reports? Not too likely. Reading through the pamphlet on prostate cancer published by *German Cancer Care* in 2009, for instance, a man will learn that according to experts, PSA tests are an important method for early detection, and that 10-year survival rates are higher than 80% (p. 15). He may also read a press release about the European randomized trial on prostate cancer screening, which states that PSA screening reduced mortality from prostate cancer by 20%. After having consulted different sources and seen various statistics, does the man now have all information to make an informed decision? No. But he may not even notice. To begin with, he may not find out that he has been misled by the 20% figure. What it refers to is a reduction from 3.7 to 3.0 in every 1000 men (age 50–69) who participate in screening, which is an absolute reduction of 0.7 in 1000, as reported in the original study (Schröder et al. 2009). Framing benefits in terms of *relative risks* (20%) is a common way to mislead the public without actually lying (see also pp. 5–9). Second, he may not know the subtle distinction between reduced cancer mortality and reduced prostate cancer mortality. The original study reported no difference in overall cancer mortality: In the screening group, 23.9 out of 1000 men died of cancer, compared to 23.8 in the non-screening group. The 0.7 out of 1000 who did not die of prostate cancer in the screening group died of another cancer. This information, however, is virtually never mentioned in health brochures, whose aim is often to increase attendance rates. Finally, chances are low that his urologist knows the scientific evidence and is able to explain to him the pros and cons of PSA screening. Only 2 out of a random sample of 20 Berlin urologists knew the benefits and harms of PSA screening (*Stiftung Warentest* 2/2004). Even if physicians know the evidence, they may practice defensive medicine in fear of



litigation and recommend the test. For instance, although only half of 250 Swiss internists believed that the advantages of regular PSA screening outweigh its harms in men older than 50 years of age, 75% recommended regular PSA screening to their patients (Steurer et al. 2009).

What to learn from this? In the context of screening, changes in survival rates have no reliable relationship to changes in mortality, due to overdiagnosis and lead time bias. The only proof that a cancer screening test saves lives comes from mortality rates, because their calculation is not affected by the way in which diagnoses are made and thus are not biased by lead time and overdiagnosis.

---

## 9 Final Remarks

Statistical illiteracy is a big obstacle for informed decision-making. Studies document that a large number of physicians do not understand cancer screening statistics and that patient pamphlets and the media report misleading and incomplete statistics. As a consequence, a large number of patients are misinformed about cancer screenings' benefits and harms. What can be done to remedy this? Every medical school should teach its students how to understand evidence in general and health statistics in particular, and statistical literacy should be assessed in continuing medical education (CME). To cultivate informed patients, elementary and high schools should start teaching the mathematics of uncertainty—statistical thinking. Guidelines about complete and transparent reporting in journals, patient 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. Informed patients will ask questions that require doctors to become better informed about medical statistics, and in turn more easily see through biased reporting and attempts to create undue hopes and fears.

In the nineteenth century, people's health improved from a combination of clean water, better hygiene, and sufficient amounts of food. The twentieth century saw the professionalization of medicine and scientific breakthroughs, but it has left us with many uninformed physicians and patients. In the twenty-first century, we need a third revolution to promote clean information for better-informed doctors and patients (Wegwarth and Gigerenzer 2014).

## References

- Berrington de González A, Mahesh M, Kim KP, Bhargavan M, Lewis R, Mettler F, Land C (2009) Projected cancer risks from computed tomographic scans performed in the United States in 2007. *JAMA Int Med* 169(22):2071–2077
- Bramwell R, West H, Salmon P (2006) Health professionals' and service users' interpretation of screening test results: experimental study. *Br Med J* 333:284–286
- Brenner JT, Hall EJ (2007) Computed tomography—an increasing source of radiation exposure. *N Engl J Med* 357:2277–2284
- Casscells W, Schoenberger A, Grayboys T (1978) Interpretation by physicians of clinical laboratory results. *N Engl J Med* 299:999–1000
- Dobbs M (2007, October 30) Rudy wrong on cancer survival chances. Retrieved July 21, 2008 from <http://blog.washingtonpost.com/>
- Eddy DM (1982) Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A (eds) *Judgment under uncertainty: heuristics and biases*. Cambridge University Press, Cambridge, pp 249–267
- Fahey T, Griffiths S, Peters TJ (1995) Evidence based purchasing: understanding results of clinical trials and systematic reviews. *Br Med J* 311(21 October):1056–1059
- Farrow L, Taylor WC, Arnold RM (1992) Absolutely relative: how research results are summarized can affect treatment decisions. *Am J Med* 92:121–124
- Ghosh AK, Ghosh K (2005) Translating evidence-based information into effective risk communication: current challenges and opportunities. *J Lab Clin Med* 145:171–180
- Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S (2007) Helping doctors and patients make sense of health statistics. *Psychol Sci Publ Interest* 8:53–96
- Gigerenzer G, Gray JAM (2011) Launching the century of the patient. In Gigerenzer G, Gray JAM (eds) *Better doctors, better patients, better decisions: envisioning healthcare 2020*. Strüngmann Forum Report. (vol 6, pp 1–19). MIT Press, Cambridge
- Gigerenzer G, Hoffrage U (1995) How to improve Bayesian reasoning without instruction: frequency formats. *Psychol Rev* 102:684–704
- 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
- Gigerenzer G, Wegwarth O, Feufel M (2010) Misleading communication of risk: editors should enforce transparent reporting in abstracts. *Br Med J* 341:791–792
- Gøtzsche P, Nielsen M (2006) Screening for breast cancer with mammography. *Cochrane Database of Syst Rev*, 4(CD001877)
- Gøtzsche P, Nielsen M (2011) Screening for breast cancer with mammography. *Cochrane Database Syst Rev*, 1(CD001877)
- Hoffrage U, Gigerenzer G (1998) Using natural frequencies to improve diagnostic inferences. *Acad Med* 73:538–540
- Hux J, Naylor C (1995) Communicating the benefits of chronic preventive therapy: does the format of efficacy data determine patient's acceptance of treatment? *Med Decis Making* 15:152–157
- Kaiser T, Ewers H, Waltering A, Beckwermert D, Jennen C, Sawicki PT (2004) Sind die Aussagen medizinischer Werbeprospekte korrekt? *Arznei-Telegramm* 35:21–23
- Labarge AS, McCaffrey RJ, Brown TA (2003) Neuropsychologists' ability to determine the predictive value of diagnostic tests. *Clin Neuropsychol* 18:165–175
- Malenka DJ, Baron JA, Johansen S, Wahrenberger JW, Ross JM (1993) The framing effect of relative and absolute risk. *J Gen Int Med* 8:543–548
- Naylor CD, Chen E, Strauss B (1992) Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Int Med* 117:916–921
- Nuovo J, Melnikow J, Chang D (2002) Reporting number needed to treat and absolute risk reduction in randomized controlled trials. *J Am Med Assoc* 287:2813–2814

- Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE (2002) Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *The Lancet* 359:909–919
- Nyström L, Larsson LG, Wall S, Rutqvist LE, Andersson I, Bjurstam N, Tabár L (1996) An overview of the Swedish randomised mammography trials: total mortality pattern and the representativity of the study cohorts. *J Med Screening* 3:85–87
- Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Auvinen A (2009) Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 360 (13):1320–1328
- Schüssler B (2005) Im Dialog: Ist Risiko überhaupt kommunizierbar, Herr Prof. Gigerenzer? *Frauenheilkunde Aktuell* 14:25–31
- Schwartz L, Woloshin S, Dvorin EL, Welch H (2006) Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks. *Br Med J* 333:1248–1252
- Schwartz LM, Woloshin S, Black WC, Welch HG (1997) The role of numeracy in understanding the benefit of screening mammography. *Ann Int Med* 127:966–972
- Schwartz LM, Woloshin S, Fowler FJ, Welch HG (2004) Enthusiasm for cancer screening in the United States. *J Am Med Assoc* 291:71–78
- 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
- Sirovich B, Welch H (2004) Cervical cancer screening among women without a cervix. *J Am Med Assoc* 291:2990–2993
- Steurer J, Held U, Schmidt M, Gigerenzer G, Tag B, Bachmann LM. (2009). Legal concerns trigger PSA testing. *J Eval Clin Prac* 15:390–392
- Thaler RH, Sunstein CR (2008) *Nudge: improving decisions about health, wealth, and happiness*. Yale University Press, New Haven
- Wegwarth O, Gaissmaier W, Gigerenzer G (2011) Deceiving numbers: survival rates and their impact on doctors' risk communication. *Med Decis Making* 31(3):386–394
- Wegwarth O, Gigerenzer G (2011) "There is nothing to worry about": Gynecologists' counseling on mammography. *Patient Educ Couns* 84(2011):251–256
- Wegwarth O, Gigerenzer G (2013) Trust-your-doctor: a simple heuristic in need of a proper social environment. In: Hertwig R, Hoffrage U, The ABC Research Group (eds) *Simple heuristics in a social world*. Oxford University Press, New York, pp. 67–102
- Wegwarth O, Gigerenzer G (2014) Improving evidence-based practices through health literacy—in reply. *JAMA Int Med* 174(8):1413–1414
- Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G (2012) Do physicians understand cancer screening statistics? A national survey of primary care physicians in the U.S. *Ann Int Med* (156):340–349
- Welch HG, Schwartz LM, Woloshin S (2000) Are increasing 5-year survival rates evidence of success against cancer? *J Am Med Assoc* 283(22):2975–2978
- Woloshin S, Schwartz LM (2012) How a charity oversells mammography. *BMJ* 345:e5132
- Young JM, Glasziou P, Ward JE (2002) General practitioners' self rating of skills in evidence based medicine: a validation study. *BMJ* 324:950–951