Erschienen in: Handbook of risk theory : epistemology, decision theory, ethics, and social implications of risk ; Band 2 / Sabine Roeser ... [et al.] (eds.). - Dordrecht [u.a.] : Springer, 2012. - S. 621-660. - ISBN 978-94-007-1432-8

Risk Communication in Health

Nicolai Bodemer · Wolfgang Gaissmaier
Max Planck Institute for Human Development, Berlin, Germany

Abstract: Policy makers, health professionals, and patients have to understand health statistics to make informed medical decisions. However, health messages often follow a persuasive rather than an informative approach and undermine the idea of informed decision making. The current practice of health risk communication is often biased: Risks are communicated one sided and in nontransparent formats. Thereby, patients are misinformed and misled. Despite the fact that the public is often described as lacking basic statistical literacy skills, statistics can be presented in a way that facilitates understanding. In this chapter, we discuss how transparent risk communication can contribute to informed patients and how transparency can be achieved. Transparency requires formats that are easy to understand and present the facts objectively. For instance, using statistical evidence instead of narrative evidence helps patients to better assess and evaluate risks. Similarly, verbal probability estimates (e.g., "probable," "rare") usually result in incorrect interpretations of the underlying risk in contrast to numerical probability estimates (e.g., "20%," "0.1"). Furthermore, we will explain and discuss four formats – relative risks, conditional probabilities, 5-year survival rates, and single-event probabilities – that often confuse people, and propose alternative formats – absolute risks, natural frequencies, annual mortality rates, and frequency statements – that increase transparency. Although research about graphs is still in its infancy, we discuss graphical visualizations as a promising tool to overcome low statistical literacy. A further challenge in risk communication is the communication of uncertainty. Evidence about medical treatments is often limited and conflicting, and the question arises how health professionals and laypeople deal with uncertainty. Finally, we propose further research to implement the concepts of transparency in risk communication.

Introduction

Understanding health statistics is one basic prerequisite for making health decisions. Policy makers evaluate health statistics when implementing health programs, insurance companies assess the cost-effectiveness of health interventions, and doctors and patients need to know the chances of harms and benefits of different treatment alternatives. The channels available to inform decision makers about risks are manifold, and so are the ways risks can be framed. A widespread phenomenon is what we call biased reporting in risk communication. By biased, we mean two things: First, information is incomplete and one sided. For instance, benefits of a health treatment are reported, while drawbacks are omitted. Second, the continuous use of nontransparent and incomprehensible risk communication formats misleads decision makers. In this chapter, we discuss the interaction between the fact that most people have difficulties with statistical information and the way health risks can be represented. The chapter is organized as follows:

1. *What constitutes "good" risk communication?* We start this chapter by discussing the objective of risk communication.
2. *Current practice in health risk communication.* We describe current drawbacks in the practice of risk communication.
3. *How good are experts and laypeople at dealing with risks and uncertainties?* We present evidence about the public's problems in adequately interpreting statistical information.
4. *The role of numbers and words in risk communication.* We discuss the role of narrative and verbal information, in comparison to statistical information.

5. *Transparent risk communication: How to overcome statistical illiteracy and innumeracy.* We present alternative formats that improve statistical comprehension, in contrast to those frequently used in practice.
6. *Further research.* We point out important directions for future research to make our society more risk literate.

What Constitutes "Good" Risk Communication?

Of central importance, but at the same time the subject of much controversy, is the issue of what the goal of risk communication ought to be. To put it differently, what is the standard by which risk communication should be evaluated? There are at least two different perspectives a communicator can adopt: one is persuasive, the other informative (or educative). Despite commonalities between these two perspectives, there is an area of tension resulting from the different objectives each of the views follows. Let us first discuss persuasion.

The press release of the first European Randomized Study of Screening for Prostate Cancer (ERSPC) stated, "Screening for prostate cancer can reduce deaths by 20%. ERSPC is the world's largest prostate cancer screening study and provides robust, independently audited evidence, for the first time, of the effect of screening on prostate cancer mortality" (Wilde 2009). This news was celebrated as a successful demonstration of the benefits of prostate-specific antigen (PSA) screening. Based on this statement, policy makers and doctors could argue for regular PSA tests, and men might express their willingness to participate in the screening program. However, the actual benefits of screening for prostate cancer with PSA tests are not as clear as they seem, as we will demonstrate later in this chapter. Similarly, health advertisements usually promote behavioral change. For example, an advertisement for screening for vascular diseases appeals with the admonition, "Don't be a victim" (see Gigerenzer et al. 2007).

More extreme attempts to change people's attitudes and intentions are fear appeals (for a meta-analysis, see Witte and Allen 2000). For instance, antitobacco campaigns show pictures of smokers' lungs or mouth cancer to demonstrate the consequences of smoking. The aim of fear appeals is less to educate the public about health interventions than to promote and encourage health behavior change. Frosch et al. (2007) evaluated health advertisements and found that the vast majority of those aired on TV made emotional appeals, and only about one-fourth gave explicit information about risk factors, prevalence, and condition causes. These campaigns are not solely run by the pharmaceutical industry, but also by health authorities and health associations. The term "social marketing" has been coined to describe the application of "marketing principles and techniques to create, communicate and deliver value in order to influence target audience behaviors that benefit society (public health, safety, the environment, and community development) as well as the target audience" (Kotler and Lee 2007, p. 7).

The underlying assumption of the persuasive approach is that people's motivation and ability to engage in health decisions is rather low in the first place and hence deviate from a "normative" standard – however that might be defined. From this point of view, the key measure for successful risk communication is behavioral change that is reflected in more favorable attitudes toward health (prevention) programs, higher intentions to participate, and finally higher attendance rates.

The alternative perspective – that is, the informative approach – begins with the assumption that people are able to take responsibility for their health and make individual and informed health decisions. A decision per se is not right or wrong – it always depends on the patient's personal preferences, values, and needs. Some patients prefer watchful waiting to invasive treatments; others prefer rapid treatment of abnormalities. Some patients accept severe side effects of treatments if the benefit is high, while others do not. For instance, it has been reported that patients are willing to accept higher risks of severe side effects than their physicians (Heesen et al. 2010). The concept of health communication as information is related to the paradigm shift from the classic notion of a paternalistic doctor–patient relationship to one of shared decision making and informed consent – a mutual, interactive process between the doctor and the patient, who jointly make health decisions (e.g., Edwards and Elwyn 2009). With this in mind, the main evaluation principles in risk communication should be transparency and (gained) knowledge. Risk communication requires comprehensible, unbiased, and complete information to educate doctors and patients and provide a basis for shared decision making. Informed decisions require facts about etiological factors, epidemiological data, treatment benefits and side effects, uncertainties, and potential costs. Without knowing the risk of developing a particular disease, the chance that a treatment will lead to success, or the risk of side effects, neither policy makers nor doctors and patients can effectively make informed health decisions.

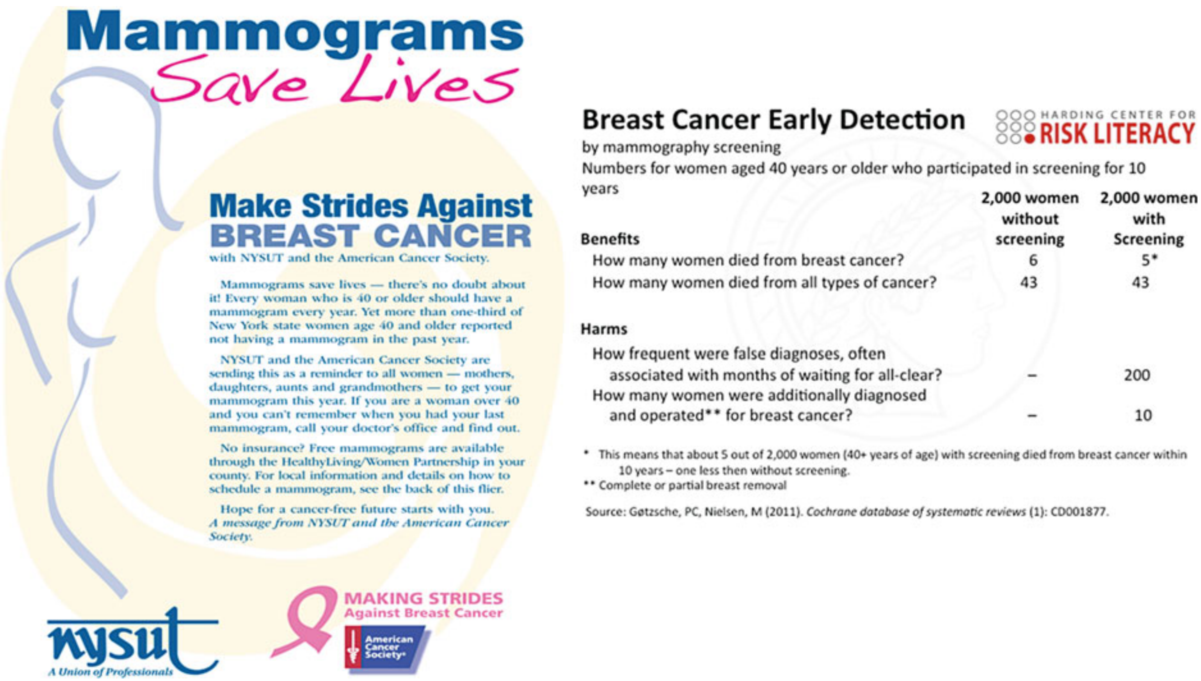
We consider ourselves proponents of the latter approach and argue that the major objective of risk communication should be informing and educating rather than persuading. However, we do not aim at discussing the "persuasion" approach.

An example contrasting the two different approaches in risk communication is given in ➔ *Fig. 24.1*. While the flyer "mammograms save lives" encourages women to participate in mammography screening and convey an illusion of certainty ("mammograms save lives – there's no doubt about it (. . .) Hope for a cancer-free future starts with you"), the facts box summarizes current scientific evidence and compares 2,000 women in a mammography group with 2,000 women not attending the screening.

Current Practice in Health Risk Communication

Consider the following fictive example: An urologist offers a 57-year-old patient a PSA test – the previously mentioned screening test to detect early stages of prostate cancer. The patient, who has never heard of this test, says to his doctor, "Well, I don't really know. What do you think I should do?" The urologist hesitates and then answers, "I think you should do the test." The patient agrees without knowing his baseline risk of prostate cancer or the benefits and harms of the PSA test. The patient trusts the doctor's recommendation and believes the doctor's decision was based on the best medical knowledge. However, this does not need to be true.

Doctors often practice what is called defensive decision making: They prescribe treatments that may not be best for their patients but that reduce their own risk of facing legal consequences. In our example, the doctor could have recommended not participating, because current scientific evidence does not show a benefit of PSA screening in the reduction of prostate cancer mortality (Djulbegovic et al. 2010; Sandblom et al. 2011). But even if the doctor did not believe in the efficacy of PSA screening, not urging the patient to have the test might cause trouble if the patient is later diagnosed with prostate cancer. Daniel Merenstein, an American urologist, informed his patient about the pros and cons of PSA screening, and the patient decided not to participate. Later, the patient developed prostate cancer and sued Merenstein, whose residency had to pay compensation of US$1 million (see Gigerenzer and Gray 2011).



Breast Cancer Early Detection
by mammography screening
Numbers for women aged 40 years or older who participated in screening for 10 years

	2,000 women without screening	2,000 women with Screening
Benefits		
How many women died from breast cancer?	6	5*
How many women died from all types of cancer?	43	43
Harms		
How frequent were false diagnoses, often associated with months of waiting for all-clear?	–	200
How many women were additionally diagnosed and operated** for breast cancer?	–	10

* This means that about 5 out of 2,000 women (40+ years of age) with screening died from breast cancer within 10 years – one less then without screening.
** Complete or partial breast removal

Source: Gøtzsche, PC, Nielsen, M (2011). *Cochrane database of systematic reviews* (1): CD001877.

Fig. 24.1
Two different ways to inform women about mammography screening. The flyer on the *left side* by the American Cancer Society (Retrieved from www.nysut.org/files/makingstrikes_070921_poster.pdf in April 2011) encourages women to participate in regular mammography screening without providing information about benefits and harms of the screening program. It states that "mammograms save lives – there's no doubt about it (. . .) Hope for a cancer-free future starts with you." The facts box on right side (Retrieved from www. http://www.harding-center.com/fact-boxes/mammography-screening in April 2011) summarizes the most important results based on the current scientific evidence and informs rather than persuades. It contrasts 2,000 women aged 40 and older who participate in mammography screening over 10 years with 2,000 of the same age who do not. Besides the benefits of the screening program, the facts box also includes information about potential harms like overtreatment

Such decisions have far-reaching consequences for doctors' behavior as well as the entire health system. For instance, many doctors in Switzerland order PSA tests for their patients but would not participate themselves (Steurer et al. 2009).

The Seven Sins in Health Care

Defensive decision making is just one of seven "sins" in health care that Gigerenzer and Gray (2011) identified. They have called for the "century of the patient" to demonstrate the importance of a radical change in health policy. This change centers on fostering patients who understand health risks and who are willing to take responsibility for their own health decisions based on transparent and unbiased information. A misinformed patient is the result of the seven sins: biased funding in medical research, biased reporting in medical journals, biased reporting in pamphlets, biased reporting in the media, conflicts of interests, defensive decision making, and last but not least, doctors' lack of fundamental health literacy skills (see ➲ *Table 24.1*). Although these sins are more or less linked to each other, we will primarily address the issues of biased reporting and the lack of statistical literacy in health professionals. Surprisingly, even doctors have trouble understanding medical evidence and are prone to being deceived by statistics, as we will demonstrate later.

Biased Reporting in the Medical Literature

To explain what is meant by biased reporting in the medical literature, let us again take the example of the press release of the European trial for PSA screening. It stated that PSA testing reduces the risk of dying from prostate cancer by 20%. What does this number mean? It means that out of every 1,410 men who regularly participated in prostate cancer screening, one less died of prostate cancer than in an equally sized group of men who did not participate (Schröder et al. 2009). Additionally, 48 of the 1,410 men were unnecessarily treated and hence subjected to potential incontinence and impotence and that overall mortality was also unaffected. Communicating risk as a 20% risk reduction or as the number of men needed to screen to save one life makes quite a difference.

Evaluations of abstracts in leading medical journals have shown that the majority of reports fail to report absolute risks in addition to relative numbers (Schwartz et al. 2006; Sedrakyan and Shih 2007; Gigerenzer et al. 2010). Another form of biased risk communication in the medical literature is mismatched framing: Benefits are presented in relative risk reduction formats and appear rather large, whereas side effects are presented in absolute terms and appear smaller. Thereby benefits are overestimated, side effects underestimated. An even more extreme way to misinform is the omission of any side effects.

Biased Reporting in the Media and Pamphlets

In contrast to presenting deceiving numbers, many health pamphlets do not present any numbers at all. A pamphlet informing the public about the human papillomavirus (HPV) vaccine – an innovative vaccine to prevent the risk of cervical cancer – states the following: "For two years,

Table 24.1

The seven sins in health care: The table summarizes the seven factors that contribute to misinformed patients identified by Gigerenzer and Gray (2011)

The seven sins	Example
Biased funding of research	Out of estimated US$160 billion spent on research and development in health in the United States, more than half was sponsored by the pharmaceutical, biotechnical, and medical technology industry (see Gigerenzer and Gray 2011)
Biased reporting in medical journals	Out of 222 articles published in leading medical journals between 2003 and 2004, 150 failed to report the underlying absolute risk in the abstract (see Schwartz et al. 2006)
Biased reporting in health pamphlets	Out of 27 pamphlets informing about breast cancer screening in Germany, only ten informed about lifetime risk of developing breast cancer, two reported about risk reduction of death from breast cancer in relative formats, two in absolute formats, and one presented number needed to treat (see Kurzenhäuser 2003)
Biased reporting in the media	Out of 202 German Web sites and newspaper reports informing the public about HPV vaccination, 116 reported information about baseline risk of developing cervical cancer (96 gave correct estimates); 102 out of the 202 reports reported about pros and cons of the vaccination in a balanced way. Correct estimates of risk reduction were provided in 14 articles only (see Bodemer et al. submitted)
Commercial conflicts of interest	After a drug has been approved, doctors are offered money for each patient they put on the drug by companies (between 10 and 1,000€ per patient in Germany). In 2008, out of 150,000 private medical practices, about 85,000 participated in such programs (see Gigerenzer and Gray 2011)
Defensive medicine	In Switzerland, 41% of the general practitioners and 43% of internists reported that they sometimes or often recommend PSA tests for legal reasons. In other words: They order a test for patient which they would not order for themselves (see Steurer et al. 2009)
Doctor's lack of understanding health statistics	Ninety-six out of 160 gynecologists overestimated the positive predictive value of mammography and 29 underestimated the value, despite the fact that all relevant information was available (see Gigerenzer et al. 2007)

young women have had the possibility to get vaccinated against HPV. Worldwide, 50 million vaccines have been administered. In Germany, the media has reported controversy about the vaccine, while doctors and scientists are convinced of the certainty and efficacy of the vaccine." What does this statement tell a young girl or her parents who are considering having her vaccinated? Does it mean that the vaccine reduces the risk of suffering from cervical cancer by 100%? Does the vaccine cause no side effects? Does the protection last a lifetime?

The case of the HPV vaccine is exemplary and of particular interest for two reasons: First, vaccination campaigns affect large parts of the population, primarily young girls between the ages of 12 and 16. Second, the HPV vaccine has prompted extensive media coverage, because some researchers have questioned whether it has been sufficiently evaluated (Dören et al. 2008). While the pamphlet conveys certainty and cites trustworthy and convinced experts, this only reflects half the story. We conducted a media analysis in two countries – Germany and Spain – to evaluate media coverage and how the public was informed about the HPV vaccine in Web sites and newspapers. Most of the media reports did not provide any information about the prevalence, etiology, efficacy, or uncertainties of the vaccine (Bodemer et al. submitted). It is clear through content analyses of other health communications, such as pamphlets about mammography (Kurzenhäuser 2003) or colon cancer screening (Steckelberg et al. 2001) as well as media reports about medications (Moynihan et al. 2000), that the media lack complete and balanced statistical information about risks, benefits, harms, and costs. Numbers are either not provided at all or are provided in nontransparent formats that mislead the public. This is alarming, since the mass media constitute the most prominent channels of communication about health innovations and treatments to the public (Grilli et al. 2009).

Consequences of Biased Reporting

Biased reporting undermines shared decision making and has consequences for the individual patient as well as for the health system. When the UK Committee on Safety for Medicine stated that the risk of life-threatening blood clots in legs or lungs is increased by 100% when using the third generation of the oral contraceptive pill, the public was appalled. As a consequence, many women stopped taking the pill, which resulted in unwanted pregnancies and abortions. But what did this 100% actually mean? Studies revealed that instead of 1 in 7,000 women who took the second generation of the contraceptive pill suffering blood clots, 2 in 7,000 who took the third generation pill did. This is equivalent to a relative increased risk of 100%, which in absolute numbers corresponds to a risk increase in 1 in 7,000 (example taken from Gigerenzer and Gray 2011).

Another example is the fact that treatment benefits are often overestimated. When women and men in nine European countries were asked to estimate the effect of PSA screening and mammography on prostate cancer and breast cancer mortality reduction, respectively, they highly overestimated the benefits. Especially those who consulted their doctors or health pamphlets were particularly prone to overestimation (Gigerenzer et al. 2009).

Another example is the consequence of false-positive test results, the fact that a test can erroneously signal a disease. False-positive tested patients often receive follow-up care despite the absence of disease, a phenomenon called overtreatment. Lafata et al. (2004) estimated incremental costs for false-positive results averaged over different screenings to be $1,024 for men and $1,171 for women, respectively, in the year following diagnosis. Moreover, besides unnecessary costs, false-positives create unwarranted anxieties and fears among patients.

These are just three examples that illustrate the dramatic consequences of biased reporting for health decisions. We will use these and other examples to better demonstrate and contrast the influence of different formats of risk communication.

How Good Are Experts and Laypeople at Dealing with Risks and Uncertainties?

The public is often described as lacking the fundamental skills to deal with numerical information. Two terms have been coined to illustrate this phenomenon: (collective) "statistical illiteracy" (Gigerenzer et al. 2007) and "innumeracy" (Paulos 1988). Both concepts refer to the widespread inability to understand quantitative information and to perform basic mathematical operations. But why is statistical literacy and numeracy so important for health decisions? Lipkus and Peters (2009) defined six main functions of numeracy that directly affect health decisions: Numeracy facilitates computation, encourages information search, improves interpretation of numerical information, facilitates the assessment of likelihood and value, can increase or decrease involvement in numerical data, and can consequently promote behavioral change.

Statistical (Il-)literacy in Health

Gigerenzer et al. (2007) defined 13 principles of minimal statistical literacy. One of the key competences is the ability to deal with uncertainty. People tend to sustain an illusion of certainty – an ignorant perspective in a world that cannot guarantee any certainty at all (Gigerenzer 2002). For instance, when people rated which of five tests (DNA, fingerprint, HIV, mammography, expert horoscope) yield absolutely certain results, the majority (78%) believed that DNA tests do so. Furthermore, 63% believed in the certainty of fingerprint and HIV-test results, 44% stated that mammography leads to certain outcomes, and 4% even believed in horoscopes (Gigerenzer et al. 2007). One might think that experts are not prone to this illusion, but the opposite is true. In an undercover study, a client, who was explicit about not belonging to a risk group, asked 20 professional AIDS counselors the following questions in the mandatory pretest counseling session: Could I possibly test positive if I do not have the virus? And if so, how often does this happen? The vast majority stated that the test could not err and that it was absolutely impossible to receive false-positive results, which is, of course, not true, even though false-positives are rare (Gigerenzer et al. 1998).

Therefore, the first step to becoming statistically literate is to abandon this illusion and accept living with uncertainty. Minimal statistical literacy in health also subsumes an understanding of basic statistical concepts such as sensitivity, specificity, transforming conditional probabilities into natural frequencies, and the possibility of false alarms in medical screening tests as well as an understanding of the magnitude of treatment effects. All these concepts will be explained in this chapter. In addition, statistical literacy encompasses a grasp of the quality of scientific evidence and potential underlying conflicts of interests in medical research. For instance, the gold standard for evaluating a medical treatment is a randomized control trial (RCT). However, for many medical treatments no RCT is available and scientific evidence is inconclusive or even conflicting. Patients need to distinguish between different qualities of medical evidence. Another crucial distinction for decision making in health addresses the perspective from which a risk is evaluated. First, imagine a woman who knows that of 100,000 like her, 15 will have cervical cancer. She might decide not to participate in pap smear screening to identify early stages of cervical dysplasia since her baseline risk is rather low. Now, imagine a health policy decision maker: The pap smear screening reduces the annual incidence of cervical cancer in Germany by a total 10,400 women (Neumeyer-Gromen et al. in press). In this case, a national program to implement pap smear screening might be appreciated. Thus, depending on which perspective is taken, the evaluation of a treatment has different implications.

The second approach to assessing people's ability to deal with mathematical concepts is numeracy. In a broader sense, numeracy is defined as "the aggregate of skills, knowledge, beliefs, dispositions, and habits of mind – as well as the general communicative and problem-solving skills – that people need in order to effectively handle real-world situations or interpretative tasks with embedded mathematical or quantifiable elements" (Gal 1995, cited in Reyna et al. 2009). A more concrete definition of health numeracy is given by Golbeck et al. (2005): "the degree to which individuals have the capacity to access, process, interpret, communicate, and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information needed to make effective health decisions." Some also subsume the ability to read and understand graphs under the term "health numeracy" (e.g., Ancker and Kaufman 2007), but we use the term "graph literacy" to define the ability to use visualizations (Galesic and Garcia-Retamero 2010). Moreover, Golbeck et al. (2005) differentiate four levels of health numeracy: Basic health numeracy encompasses the ability to identify numbers and correctly interpret quantifications. Computational health numeracy includes the ability to count and to conduct simple manipulations of numbers and quantities. Concepts of inference, estimation, proportions, frequencies, and percentages are represented on an analytical level of health numeracy. Finally, statistical health numeracy involves an understanding of biostatistics, the ability to compare numbers on different scales, and the critical analysis of risk ratios or life expectancy. Similar to statistical literacy, health numeracy also incorporates the understanding of scientific concepts, such as randomization and the double-blind study. Likewise, Reyna et al. (2009) reviewed the literature on numeracy and defined three levels of numeracy: The lowest level covers concepts of the real number line, time, measurement, and estimation. The middle level requires simple arithmetic operations and the comparison of magnitudes, while the highest level consists of an understanding of ratios, fractions, proportions, percentages, and possibilities.

Measuring Numeracy

Different measures have been developed to assess people's numeracy skills. Objective scales assess competence with items that measure basic, computational, analytical, or statistical abilities. For example, a simple three-item scale by Schwartz et al. (1997) requires the conversion of percent into proportion and vice versa and the estimation of the expected numbers of heads in 1,000 coin tosses. This scale was the basis for an 11-item numeracy scale developed by Lipkus et al. (2001). An alternative way to measure numeracy is with the Subjective Numeracy Scale, which asks subjects to indicate their confidence in their own mathematical skills and preferences for numerical versus verbal risk information (Fagerlin et al. 2007). Subjects have to rate how easily they can calculate a 15% discount on a T-shirt or whether they prefer weather forecasts that state a probability of rain (e.g., 20% chance of rain tomorrow) as opposed to a verbal description (e.g., a small chance of rain tomorrow). The advantage of the Subjective Numeracy Scale is that subjects are not tested but rather are allowed to estimate their own abilities and preferences. The scale showed satisfactory correlations with objective scales and is easy to apply (Zikmund-Fisher et al. 2007; Galesic and Garcia-Retamero 2010).

Numeracy in Experts and Laypeople

So how widespread is innumeracy? The Programme for International Student Assessment (PISA) in 2003 assessed mathematical and problem-solving skills of 15-year-olds in 24 countries. The results revealed low mathematical literacy skills in the United States and Germany – especially in concepts such as uncertainty and quantity. In 2007, the National Assessment of Educational Progress (NEAP) assessed students' mathematical performance. Only 22% of the students at grade 12 performed at a proficient level or above; 37% performed at basic level, and 41% even below basic level (Grigg et al. 2007; for an overview see Reyna et al. 2009). However, these results are not surprising since statistics and probability calculation are rarely implemented in school curricula. Nor is it surprising that adults have similar major difficulties in performing simple computations. The National Adult Literacy Survey (NALS) includes one scale measuring quantitative abilities. It demonstrated that 47% of the adults surveyed had very low quantitative literacy scores and difficulties in performing simple mathematical operations (Kirsch et al. 2007). These results were replicated by the National Assessment of Adult Literacy in which 36% of the subjects had a maximum of basic quantitative abilities (Kutner et al. 2006). Galesic and Garcia-Retamero (2010) compared numeracy skills in Germany and the United States using national probabilistic samples. On average, numeracy skills were higher in Germany (average proportion of correct items: 68.5% vs. 64.5%) with a greater difference between literate and illiterate in the United States. In other studies, even in well-educated samples only 16–25% of the subjects gave correct answers to all three items of the short numeracy scale (Lipkus et al. 2001; Schwartz et al. 1997). In general, men achieve higher scores than women, younger people higher scores than older people, and more educated people higher scores than those less educated.

Consequences of Innumeracy

A growing body of literature has revealed consequences of the lack of statistical literacy and numeracy skills. In one study, women read data about mammography screening and breast cancer mortality and assessed their personal risk of dying of breast cancer with and without screening. Women with low numeracy skills (none of the three items in the short numeracy scale answered correctly) had an accuracy rate of 5.8%; in comparison, those women with high numeracy skills (3 of 3 items correct) showed an accuracy rate of at least 40% (Schwartz et al. 1997). In another study, subjects were confronted with the baseline risk of a hypothetical disease and had to choose between two treatments. Benefits of the treatments were presented as number needed to treat, relative risk reduction, absolute risk reduction, or a combination of these formats. Independent of the format, high-numeracy subjects were more successful in identifying the more beneficial treatment and correctly calculating the effect of treatment for a given baseline risk than less numerate subjects (Sheridan et al. 2003). Low-numeracy subjects were also more prone to framing effects (Peters et al. 2006). Treatment effects can either be framed positively by stating that 80 of 100 patients survive a treatment or negatively by stating that 20 of 100 patients actually die. Differences between the two frames affect decisions, more so in less numerate subjects than in highly numerate students. In addition, less numerate people have more difficulties transforming one representation format (e.g., frequency "20 of 100") into another (e.g., probability "20%"): Whereas highly numerate people give consistent risk estimations independent of the format, less numerate people give lower risk estimates under probability than frequency formats. People low in numeracy also tend to overestimate their personal risks, which in turn has important consequences for the perception of treatment benefits and treatment decisions (Woloshin et al. 1999; Davids et al. 2004; Dieckmann et al. 2009). Finally, numeracy moderates denominator effects. People low in numeracy tend to ignore the information in the denominator, which leads to the misinterpretation of treatment effects when the sample size in the treatment and control group are unequal (Garcia-Retamero and Galesic 2009).

On a behavioral level, patients show difficulties in disease management. For instance, diabetes patients low in health literacy – the ability to perform the basic reading tasks needed to function in the health-care environment – and numeracy showed a poorer anticoagulation control (Estrada et al. 2004). Rothman et al. (2006) investigated the perception and interpretation of food labels in 200 primary care patients. Even though most patients indicated that they frequently used food labels and stated that these labels are generally easy to understand, many patients misunderstood information about serving size, misapplied extraneous material on the food label, and performed incorrect calculations.

The Role of Numbers and Words in Risk Communication

The communication of risks does not necessarily require an understanding of numerical information. Instead of relying on statistics, information about treatment benefits or harms can be based on the experiences of doctors and patients. Furthermore, verbal probability estimates describe risks without using data. In the following, we will describe the discrepancy between statistical and narrative evidence, and the influence of verbal probability estimates as opposed to numerical probability estimates on risk perception.

Narrative Versus Statistical Evidence

Imagine a woman age 53 must decide whether to participate in mammography screening. She decides to ask her doctor about the test. The doctor gives her the following information: "Here is what we know: Think about two groups of women at age 40 or older. In each group are 2,000 women. Whereas one group receives biannual mammography screening, the other group does not receive any screening. After 10 years, the breast cancer mortality in the two groups is compared. In the screening group, 5 out of 2,000 women died of breast cancer, whereas in the control group, 6 out of 2,000 died of breast cancer. Mammography screening prevented 1 breast cancer death out of 2,000 women." The woman is not convinced to participate in the screening. On her way back home, she meets her neighbor – a 62-year-old woman. She asks her whether she has ever participated in mammography screening and receives the following answer: "Oh, yes, fortunately, I did. About 6 years ago, my doctor advised me to have a mammogram. At that time, I didn't really know what it was and didn't know a lot about breast cancer either. But I thought it couldn't harm and did it. Then, the mammogram turned out to be positive. Of course, I was shocked. But the doctor told me that my chances are very good, since the cancer was detected at an early stage. I had a mastectomy, and since then, I'm doing fine. You can imagine how happy I am that I had the mammogram." After talking to her neighbor, the woman is convinced – she will make an appointment for a mammogram tomorrow.

This example illustrates two different types of evidence a decision maker is often confronted with: statistical evidence and narrative (anecdotal) evidence. Whereas the latter usually encompasses stories and experiences from single cases ($N = 1$), the former summarizes data of larger samples ($N > 1$). Which of the two types of evidence is more persuasive? Reinard (1988) reviewed the literature on statistical and narrative evidence and found little support for an advantage of statistical messages over narrative messages. Anecdotes and stories are more vivid, lively, and emotionally charged (Nisbett and Ross 1980; Taylor and Thompson 1982) or in other words, "it is generally accepted that stories are more concrete, more imagery provoking, and more colorful than statistics that are often abstract, dry, and pallid" (Baesler and Burgoon 1994). Consequently, narrative evidence increases personal relevance, especially when the receiver can identify with the narrator – as the 53-year-old woman did with her neighbor. In contrast, the statistical evidence provided by the doctor appears abstract and less imagery provoking. However, statistical evidence offers some advantages over anecdotes that are of particular importance for a decision maker. Statistical evidence provides information about baseline risks and treatment benefits based on a larger sample size. In comparison to stories, statistics are more factual, objective, and scientific and thereby establish a basis for credibility and trust (Baesler 1997). In their meta-analysis, Allen and Preiss (1997) also found a slightly more persuasive effect of statistical evidence than narrative evidence in different settings.

Both statistics and narratives are common in medical decision making. Ubel et al. (2001) asked subjects to choose between two treatments for angina – bypass surgery and balloon angioplasty. Both kinds of evidence were presented to the decision maker: Statistical evidence for bypass surgery showed a 75% success rate and balloon angioplasty a 50% success rate. When the narrative evidence was proportionate – in other words, when it reflected the statistical success rates (i.e., three pro statements and one contra statement for bypass surgery, and one pro and one contra statement for balloon angioplasty), 44% selected bypass surgery. However, when the number of narratives was disproportionate (one pro and one contra statement in both conditions, independent of the success rate), only 33% favored bypass surgery. Even though both conditions included statistical information, the proportion of narrative evidence affected treatment choice. In a second study, four testimonials were always presented, either proportionate or disproportionate. Whereas no significant difference in treatment choice was found between the proportionate and disproportionate format (37% and 34% chose bypass surgery), many more (58%) chose bypass surgery when numerical information was given without narratives.

The use of statistical or narrative evidence also influences risk perception. Subjects receiving a narrative reported higher personal risk than those who received statistical information, as well as higher intentions to get vaccinated, when confronted with a decision about vaccination against the hepatitis B virus (deWit et al. 2008).

In sum, the issue of which kind of evidence is more persuasive is still unresolved. As we pointed out at the beginning of this chapter, persuasion might not be an appropriate goal when conveying health messages. Instead, correct risk estimates as well as trust and credibility reflect more central evaluation measures. Since different evidence formats affect health decisions, it is crucial to understand how people perceive and interpret narrative and statistical information. Let us again consider the example of the woman facing the mammography screening decision. If she ignores the statistical evidence, she might erroneously assume that mammography is certain and prevents breast cancer deaths by 100%. Statistics help to objectively convey treatment benefits and harms and thereby help to inform and educate patients.

Expressing Probabilities with Words Versus Numbers

Risk information can, but does not necessarily have to include numbers. A meteorologist may predict a "10% chance of rain" or alternatively state that it is "unlikely" to rain tomorrow. Similarly, a physician can tell a patient that it is "very probable" that she will recover from the treatment, instead of stating that her "chances are 80%" (or in other words that 8 out of 10 patients recover). Both formats represent options for risk communication – but which is more transparent and informative? Words are more common in communication than numbers and therefore match people's internal representation, whereas the concept of probability emerged rather late in human history (Hacking 1975; Zimmer 1983). In addition, verbal probability expressions signal vagueness and uncertainty since words can never be as precise as numerical point estimates. At the same time, the imprecision of a verbal probability is its main flaw: People show immense variation in the interpretation of verbal probabilities (Budescu and Wallsten 1985; Brun and Teigen 1988). Brun and Teigen (1988) investigated how people interpret verbal probabilities and found high between-subject and within-subject variability in different domains. For instance, subjects assigned lower numerical estimates to verbal probabilities in a medical context in contrast to a context-free condition (see also Pepper and Prytulak 1974). One potential explanation of context dependency in the interpretation of verbal probability estimates is perceived base rate (Wallsten et al. 1986). A higher numerical probability estimate was associated with a verbal probability expression when the base rate of the event was high. This effect occurred primarily in verbal expressions of high and medium probability terms (e.g., possible, very likely), less so in low probability terms (e.g., rarely). Likewise, Weber and Hilton (1990) discussed perceived personal base rate and perceived severity as factors that influence the interpretation of verbal probabilities. Probability ratings were higher for more severe events, even when controlled for the base rate effect.

Verbal and Numerical Probabilities in Health

Verbal probability estimates are common in health, especially in doctor–patient communications. Doctors often describe risks and treatment effects with such verbal expressions as unlikely, probable, or certainly. However, what a doctor means by "probable" is not necessarily what a patient understands by the same term. When ranking eight different probability expressions, mothers showed higher interquartile ranges than doctors, meaning that the range of interpretation of a single expression was larger for laypeople than for experts (Shaw and Dear 1990). Can the implementation of standards in medical risk communication reduce this discrepancy? The European Commission (1998) established guidelines to indicate the frequency of side effects with five verbal terms, each representing a particular frequency (see *Table 24.2*). Knapp et al. (2004) compared how laypeople estimated the side effects in a verbal estimate condition and a numerical estimate condition with two different side effects of statins. One side effect, constipation, had a risk of 2.5%, which corresponds to a "common" event according to the guidelines; the other side effect, pancreatitis, had a risk of 0.4%, which is described as "rare." Subjects had to rate the likelihood that they would experience the side effect. The average estimated probability of occurrence for the common side effect constipation was 34.2% in the verbal condition and 8.1% in the numerical condition. For pancreatitis, the estimates were 18% in the verbal and 2.1% in the numerical condition. In general, patients give

Table 24.2
Verbal versus numerical probability estimates: Verbal probability estimates and their intended numerical equivalent from the European guideline on the readability of the label and package leaflet of medical products for human use (1998). When verbal probability estimates are presented without numerical information, laypeople tend to overestimate the occurrence of side effects. In other words, the verbal descriptors are interpreted differently by laypeople than intended by the guidelines (see Steckelberg et al. 2005)

Verbal probability estimate (proposed by European guidelines)	Numerical probability estimate (intended by European guidelines)	Estimated probability by laypeople (Mean [SD])
Very common	>10%	65 (24)%
Common	1–10%	45(22)%
Uncommon	0.1–1%	18(13)%
Rare	0.01–0.1%	8(8)%
Very rare	<0.01%	4(7)%

higher estimates for verbal probabilities than actually intended by the guidelines, which in turn influences risk perception and behavior (Berry et al. 2004).

Marteau et al. (2000) tested parents in their understanding of test results for prenatal diagnostics. When a numerical format for the test outcome was used, 97% interpreted the result correctly, whereas only 91% did so when verbal probabilities were given. Gurmankin et al. (2004) compared variations in risk perception when subjects in a hypothetical cancer scenario received either a verbal message only or a verbal message plus numerical information. In general, subjects overestimated their relative risk and showed very high variation in their estimates, within and between subjects.

Preferences for Verbal Versus Numerical Probabilities

Independent of how people interpret and understand verbal or numerical probability estimates, they might have a preference for one of the two formats. Mazur et al. (1999) confronted male patients with the treatment choice of either watchful waiting or surgery now in a prostate cancer scenario. The treatment effect in the surgery-now option was described as "possible," whereas side effects were presented in numerical information (i.e., 10–25% chance of total loss of bladder control after surgery). More than half of the patients (56%) preferred numerical information. Those patients who preferred numbers chose watchful waiting more often, compared to those preferring verbal risk information. Similarly, Shaw and Dear (1990) asked parents which format doctors should use to communicate risks and found that 72% felt that they understood the numerical information and 66% actually favored doctors who gave numerical estimates. In general, findings suggest that people tend to prefer probability information in numerical formats when they search for information but use verbal probabilities when they communicate risks to others (see, e.g., Erev and Cohen 1990; Wallsten et al. 1993). Potential reasons for the preference for numerical information is that people trust numerical information more and feel more comfortable and satisfied than with verbal estimates (Berry et al. 2004; Gurmankin et al. 2004). Despite this general pattern, interindividual differences exist. Some people feel uncomfortable with numbers and shrink from statistics, while others actively search for numerical information.

In sum, statistical and narrative evidence are important sources for decision makers but at the same time affect risk perceptions and decisions differently. People tend to perceive numbers as objective and credible. Verbal estimates lead to high inter- and intraindividual variation in the interpretation of risks. However, the "strength" of statistical evidence depends on two crucial factors. First, we demonstrated that people often lack statistical literacy and numeracy skills. Even if the public prefers to base decisions on statistics, can it adequately understand them? Low numeracy results in misconceptions that undermine informed decisions. The elimination of statistical illiteracy and innumeracy requires educational programs for doctors, patients, and children to establish a risk-literate society.

The second factor refers to a very different problem. The problem of risk communication is not simply in people's minds – their inability to deal with numbers – but rather in the environment – an environment that is primarily characterized by biased and nontransparent communication formats. Different representation formats exist to express the same (numerical) information, for example, frequencies (20 of 100) and percentages (20%). Some formats mislead people and lead to false expectations. Other formats are rather intuitive and make it easy for recipients to correctly assess a risk. What makes a format transparent is its ecological structure: the match of the external representation format and the mind – that is, the cognitive capacities to recognize relationships in certain representations of complex problems (Gaissmaier et al. 2007). The second part of this chapter will focus on this issue: transparent risk communication formats and how they facilitate the interpretation of numbers.

Transparent Risk Communication: How to Overcome Statistical Illiteracy and Innumeracy

The problem of biased risk communication is less in people's lack of statistical competency, but primarily in the use of nontransparent communication formats. We will present shortcomings of relative risks, conditional probabilities, 5-year survival rates, and their transparent counterparts. Additionally, we will illustrate the potential benefits of graphical representations and discuss approaches to including uncertainty in risk communication.

Relative Versus Absolute Risks

Let us refer again to the example of the UK pill scare. When the UK Committee on Safety for Medicine stated that the risk of life-threatening blood clots in lungs and legs increased by 100%, many women stopped taking the pill. The consequences were unwarranted pregnancies and abortions. Although stating a 100% increase is not incorrect, if an absolute instead of a relative format of risk increase is used (the risk increased by 1 in 7,000 women – i.e., instead of one woman, two women in 7,000 had blood clots), the risk appears to be very different.

A relative risk is the ratio of a risk in a treatment group and the risk in a control group who did not receive a treatment (or received only a placebo). The relative risk reduction is simply calculated by subtracting the relative risk from one. An absolute risk is defined by the difference in absolute magnitudes between the two groups. A third format to express the same information is the number needed to treat, that is, the number of patients who have to be treated to prevent one death (e.g., 100 people have to get vaccinated to prevent one death). In principle, the three measures can be converted into each other if the underlying risk is known. One might think these formats can be interchangeably used in risk communication – but the opposite is true. As the pill scare example demonstrates, the perception of a treatment's risk increase highly depends on the presented format.

Malenka et al. (1993) asked patients to select one of two treatments for a hypothetical disease with equivalent efficacy, side effects, and costs. The only difference was that one medication was framed in terms of relative risk reduction and the other as absolute risk reduction. The majority of patients (56.8%) selected the medication with relative numbers; about 15% were indifferent, and about the same proportion selected the medication with absolute numbers; and 13% could not decide. Similarly, Sarfati et al. (1998) showed subjects three different (fictitious) screening programs, each in a different format – relative risk reduction, absolute risk reduction, or number needed to treat. Depending on the format, subjects' willingness to participate differed substantially. When framed as relative risk reduction, 80% intended to participate, in comparison to only 53% and 43% who did so in the absolute risk reduction and number needed to treat condition, respectively. Additionally, relative risk reduction formats lead to higher deviations in treatment decisions from expected-utility theory assumptions than absolute formats (Hembroff et al. 2004). Does the same hold true when the subjects are medical experts? Naylor et al. (1992) showed that information in the form of relative risk reduction (relative decrease of 34%) led to higher perception of treatment effects in doctors compared with absolute risk reduction (decrease from 3.9% to 2.5%) or number needed to treat (77 persons have to be treated to save one patient). Likewise, doctors' mean ratings for effectiveness of a drug that lowers cholesterol concentration depended on whether relative or absolute risk reductions were presented (Bucher et al. 1994). A diabetes prevention intervention was rated as important or very important by 86% of health professionals under a relative risk format condition, whereas only 39% gave these ratings in an absolute risk reduction condition (Mühlhauser et al. 2006).

Consistent with other reviews (e.g., Edwards et al. 2001; Moxey et al. 2003) a meta-analysis by Covey (2007) supported the conclusion that both laypeople and experts are sensitive to the way risk reduction is framed: People perceive higher treatment effects when relative risk reduction formats are used in comparison to absolute risk reduction and number needed to treat.

As previously mentioned, health reporting is most biased when different formats are used for different effects, known as mismatched framing. Describing treatment benefits in relative numbers and treatment harms in absolute numbers confuses and misleads patients. For instance, a German pamphlet about hormone replacement therapy (HRT) states the following: 60 out of 1,000 women develop breast cancer in their lives. After HRT for 10 years, 66 out of 1,000 women develop breast cancer – the absolute increase is 6 in 1,000. At the same time, only half as many of the women who take HRT develop colon cancer, compared to those who do not take HRT; in other words, HRT reduces the risk of developing colon cancer by 50%. By using two different formats to describe benefits and harms of HRT, the consumer is misled and overestimates the benefits in contrast to the harms.

However, some argue in favor of the use of relative risks and odds ratio, especially in meta-analyses. The rationale is that both formats are supposedly more stable across different subpopulations than absolute risks (e.g., Smeeth et al. 1999). In any case, this does not have any implications for risk communication, which should always be based on absolute numbers.

Lessons learned: Findings demonstrate that no relative risks should be used in risk communication. Risk reduction or risk increase ought to be presented in absolute numbers only.

Conditional Probabilities Versus Natural Frequencies

We already mentioned the illusion of certainty and the problem that patients and health professionals often believe in the certainty of medical test results. For instance, 44% in one study stated that the result of a mammogram is absolutely certain (Gigerenzer et al. 2007). But what is actually the probability that a woman has breast cancer given a positive mammogram? To illustrate what a positive mammogram means, look at the following information (Eddy 1982; see Gigerenzer et al. 2007):

– The probability of breast cancer is 1% for a woman at age 40 who participates in routine screening (this is the prevalence or base rate).
– If a woman has breast cancer, the probability is 90% that she will have a positive mammogram (this is the sensitivity or hit rate).
– If a woman does not have breast cancer, the probability is 9% that she will also have a positive mammogram (this is the false-positive rate).

The task is to estimate the probability that a woman at age 40 who had a positive mammogram actually has breast cancer. What is the correct answer? When Eddy (1982) presented a similar scenario to staff at the Harvard Medical School, 95 of 100 physicians gave an answer between 70% and 80%, though the correct answer is about 10% – or in other words, of ten women who had a positive mammogram, only about one actually has breast cancer. Why do people have problems solving this and similar tasks?

The simple answer would be that humans are not "Bayesian" and hence are not capable of calculating posterior probabilities $P(\mathrm{H}|\mathrm{D})$ based on the prior probability $P(\mathrm{H})$, the likelihood $P(\mathrm{D}|\mathrm{H})$, and the probability $P(\mathrm{D}|{-\mathrm{H}})$. In the mammography example, $P(\mathrm{D}|\mathrm{H})$ is the sensitivity (90%), $P(\mathrm{H})$ is the base rate (1%), and $P(\mathrm{D}|{-\mathrm{H}})$ is the false-positive rate (9%). The computation of the posterior probabilities requires Bayes's theorem:

$$P(\mathrm{H}|\mathrm{D}) = \frac{P(\mathrm{D}|\mathrm{H}) \cdot P(\mathrm{H})}{P(\mathrm{D}|\mathrm{H}) \cdot P(\mathrm{H}) + P(\mathrm{D}|{-\mathrm{H}}) \cdot P({-\mathrm{H}})} \tag{1}$$

Kahneman and Tversky (1972) stated that humans cannot perform Bayesian reasoning and lapse into cognitive biases. For instance, humans tend to ignore base rates when calculating conditional probabilities. As a consequence of these biases, people's judgments are often inconsistent with normative "Bayesian" prescriptions (Casscells et al. 1978; Eddy 1982). Erroneously, humans do not differentiate between the probability of a disease given a positive test result (the posterior probability), and the probability of having a positive test result given the disease (sensitivity). Bayes's theorem is the common formula in most medical and statistical textbooks to calculate posterior probabilities, but still people who should be familiar with the formula seem to have difficulties in its application.

Gigerenzer and Hoffrage (1995) challenged the assumption that people cannot solve Bayesian tasks and proposed an alternative representation format that facilitates the computational process: natural frequencies. Think again about the mammography example, but this time, the following information is given:

– Ten of 1,000 women at age 40 who participate in mammography screening have breast cancer (prevalence or base rate).
– Of these ten women, nine have a positive mammogram (sensitivity or hit rate).
– Of the 990 women who do not have breast cancer, about 89 will have a positive mammogram nonetheless (false-positive rate).

Now imagine a representative sample of 1,000 women aged 40 who participate in breast cancer screening. How many of these women with a positive test result actually have breast cancer? Of course, the answer is the same: About 1 of 10. Nevertheless, to arrive at the correct solution does not require Bayes's theorem. Instead, the calculation is much simpler: Of 1,000 women, 98 will have a positive mammogram (9 of the 10 women who actually have breast cancer – referred as a in the formula and 89 of the 990 healthy women, referred as b in the formula). Of this 98 with a positive test result, only 9 actually have breast cancer, which results in 9.2%, or about 10%.

$$P(\mathrm{H}|\mathrm{D}) = \frac{a}{a+b} \tag{2}$$

When Gigerenzer and Hoffrage presented the mammography problem in natural frequencies instead of conditional probabilities, about half of the subjects gave the correct solution in comparison to only one-quarter in the probability condition. Since then, many researchers have replicated the results. Cosmides and Tooby (1996) conducted a series of experiments and supported the hypothesis that natural frequencies lead to higher proportions of correct inferences compared to probability formats. For example, they replicated Casscells et al. (1978) study: Only 12% of their subjects arrived at the correct result when confronted with probabilities, but between 56% and 76% did so when confronted with natural frequencies.

Nonetheless, the concept of natural frequencies has aroused controversy about whether and why it facilitates Bayesian reasoning. Some researchers have argued that frequencies per se do not improve people's performances in Bayesian tasks. However, they confused natural frequencies with other kinds of frequencies (for an overview see Hoffrage et al. 2002). For instance, Macchi and Mosconi (1998) demonstrated that not all kinds of frequencies facilitated Bayesian reasoning, and Lewis and Keren (1999) reached a similar conclusion. Gigerenzer and Hoffrage (1995) stated in the original paper that the computational simplification can be obtained only for natural frequencies, not for normalized frequencies, which – just like probabilities – require Bayes's theorem. Further misunderstandings have resulted from proposed alternative explanations for the same phenomenon, for example, that the facilitating effect is based on a "nested-set structure" or on "partitive representations" (Barbey and Sloman 2007), which actually just restate the original argument.

Barton et al. (2007) proposed a statistical taxonomy subsuming three orthogonal dimensions to reduce confusion: First, the information can refer to one event only (single-event probabilities) or a set of events (frequencies). Second, different numerical representations, such as percentages, fractions, real numbers between 0 and 1, and pairs of integers, are differentiated. Third, the information can be presented in normalized formats or nonnormalized (also called conjunctive) formats. Due to the orthogonality of the dimensions, any combination is possible. For instance, expressing the mammography information in chances leads to the same computational effect as doing so with natural frequencies but refers to a single individual (Brase 2009).

In summary, findings support the conclusion that natural frequencies help people solve Bayesian tasks and understand positive predictive values. Doctors and patients can easily learn what a positive test result means and how prevalence, sensitivity, and false-positives interact. Teaching natural frequencies is also rather simple. Instead of Bayes's rule, which invites learners to forget the actual components of the formula, the principle of natural frequencies is easy to grasp and helps people convert probabilities into natural frequencies. Even children benefit from this representation format and can perform Bayesian tasks (Zhu and Gigerenzer 2006).

We already mentioned the consequences of not understanding positive predictive values in the ➲ Introduction. Imagine a woman aged 56 who has a positive mammogram. Besides being extremely worried and anxious after receiving this result, she has to undergo further examinations and treatments. However, the chances of her result just being a false-positive are 9 out of 10. The terms “overdiagnosis” and “overtreatment” have been coined to call attention to the phenomenon that many people who have a positive screening test result are actually treated, despite the absence of the disease.

Lesson learned: While people have difficulties interpreting and calculating conditional probabilities, natural frequencies facilitate Bayesian reasoning.

Single-Event Probabilities Versus Frequencies

Another advantage of frequency statements is that they always include a reference class. This is not the case for single-event probabilities. A single-event probability is defined as “a probability that refers to an individual event or person” (Gigerenzer et al. 2007). Thus, no reference classes are included, which often leads to misconceptions between a communicator and a receiver. A meteorologist forecasts that the probability of rain tomorrow is 30%. This prediction leaves room for different interpretations. It could mean that it will rain 30% of the time, in 30% of the area, or on 30% of the days like the one tomorrow (Gigerenzer et al. 2005). While the latter interpretation is correct, most people believe the other two options to be true. Similarly, stating that the probability of developing sexual problems as a consequence of a drug is 30% leaves the patient alone to his or her subjective interpretation. Again, the probability could refer to 30 out of 100 sexual encounters of a single person or to 30 out of 100 patients taking the drug. Frequency statements always include a reference class and thereby eliminate misunderstandings.

Lesson learned: Always provide the reference class to which a probability refers.

Five-Year Survival Rates Versus Mortality Rates

When evaluating a health treatment, the first question that comes to mind is whether it saves lives in the long run. Cancer screenings aim at identifying a cancer at an early stage, even before first symptoms occur. Thus cancer screenings usually increase incidence rates – the number of cancers in a given population within a given time frame. This fact prevents us from drawing conclusions about a screening's effects on life expectancy.

Probably the most common unit mentioned when evaluating health treatments is the so-called 5-year survival rate. Survival rates can be defined as the number of patients alive at a specified time following diagnosis (such as after 5 years) divided by the number of patients diagnosed (Gigerenzer et al. 2007).

$$\text{5}-\text{year survival} = \frac{\text{Number of persons diagonosed with a specific cancer still alive 5 years after diagnosis}}{\text{Number of persons diagonosed with a specific cancer in the study population}} \tag{3}$$

For example, if a screening detects 100 people who have a positive diagnosis, and 80 of them are still alive after 5 years, the 5-year survival rate is 80%. If only 20 are still alive, the 5-year survival rate is 20%. One might expect that the higher the 5-year survival rate, the better. However, there is an alternative to 5-year survival rates: (annual) mortality rates. And even more surprising, the correlation between changes in 5-year survival rates and changes in mortality rates over time is zero (Welch et al. 2000). The mortality rate is defined as the number of people in a group who die annually from a disease, divided by the total number of people in the group.

$$\text{annual mortality rate} = \frac{\text{number of persons who die from a specific cancer over 1 year}}{\text{number of persons in the study population}} \tag{4}$$

By comparing the two formulas, one thing is apparent: The 5-year survival rate only includes people who are *diagnosed* with disease in the denominator. This fact makes it prone to biases. An annual mortality rate, instead, includes the entire population at risk in the denominator. However, 5-year survival rates are the rule rather than the exception in current risk communication.

Lead-Time Bias

The first shortcoming of 5-year survival rates is the so-called lead-time bias: A higher proportion of people alive 5 years after screening does not necessarily mean that people actually live longer – it might only be an illusionary extension of life (see *Fig. 24.2a*). Assume that the 5-year survival rate for a specific cancer was 1% between 1960 and 1965. For the time between 2005 and 2010, the proportion of people alive after 5 years was 80%. The cancer in the 1960s was diagnosed when patients first showed symptoms. In the years 2005–2010, a special screening test was applied for earlier detection, even before first symptoms occurred. Thus the time to diagnosis was reduced. However, this does not necessarily mean that it prolongs life, since patients between 1960 and 1965 and 2005 and 2010, respectively, might die the same number of years after developing the cancer. The difference is that those who lived between 2005 and 2010 knew earlier; in other words, they lived longer with the diagnosis, but in fact their life expectancy was the same.

Overdiagnosis Bias

The second shortcoming is the so-called overdiagnosis bias. Not every tumor or dysplasia (that is detected) is necessarily fatal. Some tumors would have never been detected without screening, because they would never have caused any symptoms, would have resolved spontaneously,

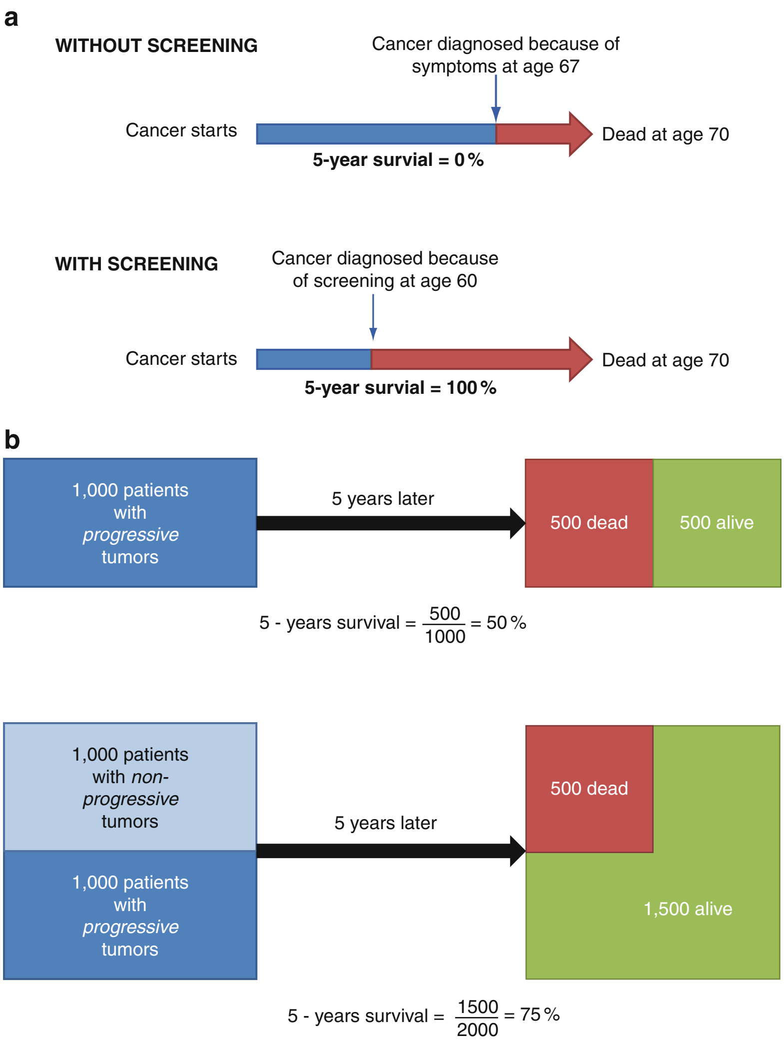


Fig. 24.2

Shortcomings of 5-year survival rates: The figure illustrates the two potential biases of 5-year survival rates (modified from Gigerenzer et al. 2007). (a) *Lead-time bias*: The arrows illustrate the course from the beginning of a disease to death. In the group without screening, cancer is diagnosed at age 67, in the screening group at age 60. However, in both groups, patients die at the same age (age 70). Whereas in the non-screening group the 5-year survival rate is 0%, it is 100% in the screening group. (b) *Overdiagnosis bias*: (1) A group of 1,000 patients with progressive tumors is monitored over 5 years. After 5 years, 500 are still alive; the survival rate is 50%. (2) The same group of 1,000 patients with progressive tumors is monitored over 5 years. Additionally, the screening detects patients with nonprogressive, indolent tumors. Again, after 5 years 500 patients died (500 out of 1,000 with progressive cancer). However, in the calculation of the 5-year survival rate, those 1,000 with nonprogressive tumors are also included hence the 5-year survival rate is 75%

or would not have been detected before the patient died of other causes. To illustrate this example, take a look at the formula 3 for the 5-year survival rates. Assume that we have 1,000 people in a screening with a positive diagnosis of having a progressive tumor based on symptoms. These 1,000 patients form the denominator to calculate the survival rate. After 5 years, half of them are still alive. Hence, the survival rate is 50%. Now imagine that our screening also detects very small and indolent tumors. These tumors are nonprogressive and hence not lethal. To the 1,000 patients with progressive tumors in the denominator, add the 1,000 patients with indolent tumors. However, the number of deaths due to the cancer is still 500. Now the survival rate is 75%. Although the number of deaths remains the same, the 5-year survival rate provides a much more favorable picture – a bias that does not affect annual mortality rates (see ➋ *Fig. 24.2b*). One well-known example is the Mayo Lung Project of the 1970s and 1980s. Smokers were assigned to either a screening group or a control group receiving no screening. Whereas 206 lung tumors were detected in the screening group, only 160 were detected in the control group. However, the overall mortality in both groups was the same. A follow-up in 1999 with patients of both groups who were still alive and had no positive diagnosis in 1983 showed that 585 patients of the screening group compared to 500 in the control group had lung cancer. One interpretation of the data is that the screening detected small and indolent tumors that were not lethal and therefore did not need any treatment (Marcus et al. 2006).

In summary, an increase in 5-year survival rates can be observed under three conditions: First, due to lead-time bias, the tumor is detected earlier, but patients still die at the same age as without screening. Second, due to overdiagnosis, more tumors are detected, but among them are indolent ones – the number of deaths is not affected. Third, screening allows for earlier detection and better treatment – in this case the screening is indeed beneficial. Mortality rates also capture the latter effect but do not fall into the traps of lead-time and overdiagnosis bias. As mentioned above, the correlation between changes in 5-year survival rates and changes in mortality rates is zero. Furthermore, a comparison of survival and mortality rates for 20 tumors between 1950–1954 and 1989–1995 showed an absolute increase in 5-year survival rates of between 3% and 50% over 5 years, whereas changes in mortality rates ranged from −80% to 259% (Welch et al. 2000). As mentioned above, those changes were completely uncorrelated.

The influence on treatment evaluations of 5-year survival rates in comparison to mortality rates was shown by Wegwarth et al. (2011). They presented physicians with either a 5-year survival rate only, annual disease-specific mortality only, or a combination of the two formats with or without incidence rates, and all the numbers were based on the same, real data. When only 5-year survival rates were presented, 66% of the physicians recommended screening; 78% were convinced of the screening's efficacy and showed the highest overestimation of the screening's benefit. In contrast, when confronted with mortality rates, only 8% gave a recommendation and only 5% considered the screening to be efficient. The two combined versions produced results in between these values. This study illustrates the influence of the format used to describe screening effects on measures of risk reduction as well as behavioral measures. Only annual mortality rates convey a transparent and unbiased picture of actual changes in mortality that could be the result from screening. As the 5-year survival rate is a highly specific medical concept, it is not surprising that patients would be unaware of its misleading potential, but that even doctors are not aware of these problems is worrying.

Lessons learned: Five-year survival rates do not allow us to adequately evaluate health interventions, particularly screenings. In contrast, annual mortality should be used to illustrate effects.

Graphical Representations

A promising alternative way to present numerical estimates is with graphs. As the saying goes, a picture is worth a thousand words. In the eighteenth century, William Playfair was the first to use bar charts and pie charts to illustrate economic and political data. Later, at the beginning of the twentieth century, the philosopher and economist Otto Neurath proposed symbols to display statistical information. Since then, different formats have been identified and used to communicate risks. The advantages of graphs are manifold (see Lipkus and Hollands 1999; Ancker et al. 2006; Lipkus 2007; Zikmund-Fisher et al. 2008a, b):

1. *Graphs often attract and maintain attention*: Attention and the expectation of successfully handling quantitative information are important prerequisites for a motivated patient to get involved in personal health decisions.
2. *Graphs foster automatic and intuitive processes*: A transparent and well-designed graph reduces the cognitive effort needed to extract and understand the information. For instance, certain formats do some of the mathematical operations for the observer and consequently facilitate understanding (e.g., part-to-whole relationships; ratio concepts).
3. *Data patterns can be detected*: Some graphs help display data over a longer period of time and show patterns and trends in the long run (e.g., lifetime incidence).
4. *Graphs can help communicate uncertainty*: A main challenge in risk communication is the inclusion of uncertainty parameters. Graphs can facilitate displaying uncertainty transparently.
5. *Graphs can overcome low numeracy*: Statistical illiteracy and innumeracy are widespread phenomena. Especially those low in numeracy tend to misunderstand risk information. Graphs offer an alternative format to help people with low numeracy understand risks and make informed decisions.

Although graphs facilitate the communication of statistical information, a variety of visualization formats can be used to provide the same information – and not all are equally effective. Again it is a question of ecologically structured information. Just like numbers, some graphs have the potential to display information in a biased way and consequently mislead the public. Additionally, understanding graphs requires basic skills to extract the relevant information and read beyond the data given – an ability described as graph literacy.

The Example of Icon Arrays

One format to visually display risks is icon arrays (pictographs). Icon arrays are graphical visualizations consisting of icons (faces, circles, figures) that represent individuals belonging to a certain group or a group as a whole (see *Fig. 24.3*).

Fagerlin et al. (2005) tested the impact of pictographs on a hypothetical treatment choice in an angina scenario. Subjects had to choose between two treatments and received anecdotal evidence about success and failure rates of the respective treatments. Anecdotal evidence was either representative of the success rates or not. The treatment with the higher success rate was chosen more often when anecdotal evidence was representative in comparison to when it was not. However, when pictographs were included, the representativeness effect of the anecdotes diminished. By including icon arrays, the influence of narrative evidence can be reduced.

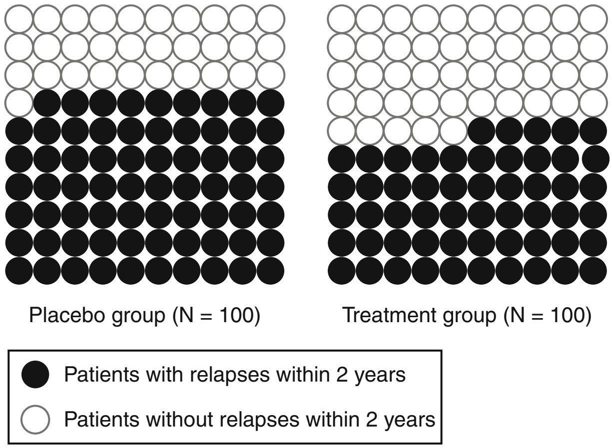


Fig. 24.3
Icon array: Benefit of interferon therapy for multiple sclerosis patients (modified from Heesen et al. 2008, p. 40). The treatment aims to reduce multiple sclerosis relapses. Displayed are two groups – a placebo group (*left*) and a treatment group (*right*) – with 100 patients in each. In the placebo group, out of 100 patients 31 patients had no multiple sclerosis relapse within 2 years. In the treatment (interferon therapy) group, 45 had no multiple sclerosis relapse within 2 years. In other words, 14 out of 100 patients benefit from interferon therapy

Icon arrays also help reduce base rate neglect (Garcia-Retamero and Galesic 2009; Garcia-Retamero et al. 2010). Subjects had to compare a treatment group with a control group. When the two groups have equal sample sizes, subjects showed no difficulties in estimating treatment effects. However, the proportion of correct estimates decreased when the groups differed in sample size – the denominator information was often neglected. Icon arrays call attention to the different sample size in the treatment and control group and thereby help to reduce base rate neglect.

It could be argued whether icon arrays facilitate understanding of risk information in addition to transparent numerical formats. But in fact, graphs have an incremental effect on risk perception. In one study, subjects received treatment effects framed as absolute or relative risk reduction with or without icon arrays (Galesic et al. 2009). Icon arrays improved performance for both framing formats. Especially those low in numeracy benefited from the additional visualization.

When Stone et al. (1997) compared a numerical risk reduction format (30–15) with stick figures to understand people's willingness to pay for improved toothpaste versus standard toothpaste, they observed a higher willingness to pay in the graphical condition. They interpreted the result as meaning that graphs led to higher risk avoidance in comparison to numbers. However, while the base rate was presented in the numerical condition, the stick figures did not represent the part–whole relationship. In later experiments (Stone et al. 2003), the stick figures were replaced by charts and stacked bar charts. The inclusion of the part–whole relationship in the graphical formats reversed the original effect: Willingness to pay was lower in the graph conditions than in the numerical conditions. Similarly, Garcia-Retamero and

Galesic (2010) found that bar charts and icon arrays primarily improve risk understanding when the population at risk is included. Pie charts and horizontal bar charts helped people with low numeracy to overcome framing effects.

Graph Literacy

Despite graphs being more intuitive and facilitating the interpretation of statistical information, their understanding requires basic skills. Galesic and Garcia-Retamero (2011) developed a 13-item scale to measure graph literacy in the medical domain. The scale is meant to differentiate people's ability to read graphical information (e.g., find the correct data), to read between the data (find relationships, e.g., to identify which proportion in a pie chart is larger), and also beyond the data (e.g., inferences and predictions that can be derived from the data). The scale covers different graph formats such as bar graphs, pie charts, line graphs, and icon arrays. The authors validated the scale on probabilistic samples in Germany and the United States and reported correlations between graph literacy and education in Germany (0.29) and the United States (0.54) and between graph literacy and numeracy in Germany (0.47) and the United States (0.55). The majority of people who are low in numeracy are also low in graph literacy and those high in numeracy are usually higher in graph literacy. Primarily those low in numeracy but high in graph literacy benefited from the inclusion of visualizations; those high in numeracy but low in graph literacy had no incremental benefits.

Notwithstanding the advantages of graphs to communicate risks, shortcomings of visual displays cannot be overlooked. People who lack basic graph literacy skills have difficulties extracting and interpreting the relevant information. Up to now, little has been known about how graphical, text, and numerical information interplay. People might shift all their attention to graphs, thereby ignoring any other information relevant to a decision. Additionally, more research is needed to know under what condition a particular visualization format works best. Some graphs are better than others for communicating lifetime prevalence or contrasting different treatment effects. Last but not least, many graphs in newspapers and the scientific literature look fancy but are in fact uninformative or even misleading. Wainer (1984) summarized the 12 most powerful techniques for displaying data badly. In sum, as Lipkus (2007) pointed out, "knowledge of how graphical displays affect risk perceptions is still in its infancy and remains, with few exceptions, a largely atheoretical research area" (p. 702).

Lessons learned: Graphs are a promising tool for improving people's interpretation of risks. However, research about graphs is still in its infancy.

Uncertainty Communication

The French philosopher Voltaire once said "Doubt is not a pleasant condition, but certainty is absurd." Our world is fundamentally uncertain and yet many people cling to the illusion of certainty. Knight (1921) previously distinguished between risk, which can be computed numerically, and uncertainty, which is immeasurable. A less strict definition of uncertainty allows estimating parameters of uncertainty probabilities, such as standard deviations, confidence intervals, or experts' confidence ratings. A further distinction is made between uncertainty and variability (Thompson 2002). Variability refers to the fact that different individuals or groups in a population have different risks. This could be due to differences in age, gender, region, or exposure to risk factors. This is different from uncertainty, which refers to imperfect knowledge. Although often ignored, uncertainties play a major role in medical decision making (see, e.g., Politi et al. 2007). First, scientific evidence is limited. Even randomized controlled trials – often regarded as the gold standard in medical research – have limitations due to design principles, sample size, and lack of validity and reliability of measures. Second, risk estimates are based on population data and therefore cannot be applied one-to-one to individuals (see example of individual vs. public health perspective on pap smear screening above). Third, risk estimates are based on past events. Their application to the present and future rests on the assumption that the environment and underlying forces do not change.

Fear of Disclosing Uncertainty

Uncertainty communication is more the exception than the rule. When in 2009 the World Health Organization proclaimed the H1N1 pandemic, it was assumed that worldwide about two billion people could become infected and between 2 and 7.4 million people could die. This forecast influenced many policy decisions, such as the implementation of H1N1 vaccination programs (Feufel et al. 2010). However, at that point of time, little was known about the actual severity and spread of the virus. In hindsight, this prediction appears exaggerated. Why do experts shrink from disclosing and communicating uncertainty?

First, many experts believe that the public is incapable of understanding uncertainty (Frewer et al. 2003). Communicating uncertainty confuses people who are mainly ambiguity averse and uncertainty intolerant (Epstein 1999). Second, experts might fear losing trust and perceived competence if they reveal that some aspects of an issue are unknown. However, as the H1N1 example illustrates, the opposite can happen: Maintaining an illusion of certainty that in hindsight turns out to be false decreases trust in the expert (Holmes et al. 2009; Feufel et al. 2010). Narrative evidence underlines this conclusion. When the Bank of England started to publish the protocols of their board meetings and transparently reveal related uncertainty in their prognoses about economic growth, the British public rated it as the most trusted institution (Gigerenzer 2007, p. 215).

How to Communicate Uncertainty

The core question is still how to transparently communicate uncertainty. Ibrekk and Morgan (1987) studied laypeople's understanding of nine different graphical uncertainty representations. Subjects received graphs of weather forecasts about the probability of snow (without any explanations) and the prediction of water depth in a flood (with explanations). As a dependent measure, subjects had to assess the most likely estimate as well as the range. Findings showed that a point estimate including a 95% confidence interval and a Tukey box were easiest to understand. The presence of an explanation led to a slight improvement. Subjects were most sure about their estimates if the point estimate plus confidence interval or a histogram was provided. Cumulative density functions and pie charts seemed improper for communicating risks. This was among the first studies investigating tools to communicate uncertainty to laypeople.

Johnson and Slovic (1995) conducted four studies to understand the influence of uncertainty information. They concluded that people are unfamiliar with uncertainty information and that the recognition of uncertainty caused trouble. However, graphical tools can help communicate uncertainty; the disclosure of uncertainty also signaled honesty for some people but was a sign of incompetence for others. To better understand laypeople's perception of uncertainty, Schapira et al. (2001) formed focus groups and asked women whether an uncertainty statement (in this case a confidence interval) should be included in risk communication. More highly educated women appreciated the inclusion of confidence intervals and interpreted the information as more complete, whereas less educated women reacted with a decrease in trust and the dilution of the actual treatment benefit. Another study investigated effects of doctor's uncertainty disclosure on breast cancer patients (Politi et al. 2010). Although women's breast cancer treatment choice and consistency with expert's opinion was independent of the doctor's disclosure of uncertainty, uncertainty communication reduced decision satisfaction.

With respect to risk perception, it is assumed that increased uncertainty leads to an increase in perceived risks. Put differently, the more uncertain a hazard, the greater the worry that is associated with it, which in turn shifts people's focus to bad outcomes (Einhorn and Hogarth 1985; Viscusi et al. 1991). Yet Kuhn (2000) found that the communication of uncertainty interacts with prior attitude. Subjects were split into two groups defined by high or low environmental concern and received five different scenarios describing environmental hazards. When the risk information was presented as a point estimate, environmental attitudes predicted environmental risk perception. However, the differences between the two groups reduced when an uncertainty statement was included, primarily because those with high environmental concern showed lower perceived risks. A potential explanation is that people with high environmental concern appreciated the uncertainty information (either as a verbal or numerical statement) and perceived the communicator as more honest – similar to the conclusion drawn by Johnson and Slovic (1995). These results are also in line with other proposals, namely, that one way to increase credibility and trust is to present uncertainty instead of maintaining an illusion of certainty (Frewer 1999; van Dijk et al. 2008).

Interestingly, uncertainty communication can improve decisions with respect to a normative criterion (Nadav-Greenberg and Joslyn 2009). Subjects played a road treatment task and had to decide whether to salt roads based on either a point estimate ("It will be 1.7°C tomorrow") or an uncertainty forecast ("It will be 1.7°C tomorrow, but there is an 18% chance that it will freeze"). For every treatment of the road, an amount of $1,000 had to be paid. However, if the decision maker did not salt the roads and the temperature dropped to freezing, a penalty of $6,000 had to be paid. According to the expected value, the roads should be treated when the probability of freezing is above 16.7%. Subjects with a deterministic forecast showed a larger deviation from expected value compared to subjects with probabilistic forecasts, resulting in a higher end budget for those who had access to the deterministic forecast *and* probability of a freeze.

In sum, evidence suggests that uncertainty communication is not as prejudicial as often believed. Laypeople understand and even appreciate uncertainty information. There is no evidence that confirms experts' fear of losing trust and perceived credibility when disclosing uncertainty – the opposite effect might be true. However, this is only a tentative conclusion; more research is needed to better understand how uncertainty communication affects and interacts with trust, credibility, and decision-making processes. People have different expectations in different domains that affect their willingness to accept uncertainty – these inter- and intraindividual differences should be examined to improve risk communication and make uncertainties more salient.

Lessons learned: Uncertainty communication is not bad per se. How it finally affects decision quality, trust, credibility, and satisfaction still requires further examination.

Further Research

Throughout this chapter, we have presented and discussed the current state of research in health risk communication. The research presented here so far has sought to support the idea that transparent risk communication can inform the public and become the basis for shared decision making. In the following, we will discuss what boundaries exist that undermine transparent risk communication and how to overcome them. The following three aspects are at the center of our discussion:

1. Research gaps: Where is further research needed?
2. Political and societal obstacles: Why is risk communication rarely put into practice?
3. Teaching statistical literacy: How can society become statistically literate?

Whereas the first point primarily addresses the research community, the latter two require transferring research into practice and demand the interaction between researchers and policy makers.

Research Gaps

While research in risk communication has already identified multiple formats and alternatives for representing statistical information, there are still many open research questions to be addressed. Here are three potential research fields.

Individual Differences

It would be interesting to look at the role of interindividual differences in understanding risk communication. People do not only differ with respect to numeracy and graph literacy skills, which have important implications for their understanding of risks. Other frequently discussed factors are age, education, socioeconomic status, intelligence, need for cognition, prior experience, and media competency. Especially the development of patient support technologies, which aim at providing tailored information to individuals, requires knowledge about how individual differences affect information search and decision making, and how to discover these differences. With the help of tailored information, patients can evaluate and select those treatments that fit their personal preferences and needs.

Let us illustrate this with a fictitious example: Imagine three patients who are thinking about participating in colon cancer screening. They have four alternatives: fecal-occult blood test, DNA test, colonoscopy, and sigmoidoscopy. The screening programs differ on various dimensions like how well the test detects early stages of cancer, how often the test errs, how invasive the treatment is, what the side effects are, and how much the test costs. One patient searches for information about all four alternatives on all dimensions. He weights each question according to his preferences and adds up his evaluations of all treatments. Finally, he selects the one with the highest score. This is called a weighting and adding strategy. However, the second patient follows a different strategy: She thinks that the test should be as good as possible in detecting early stages of cancer. Therefore, she ignores the other questions. If two tests are equally effective, she compares those alternatives on the level of side effects and chooses the one with less severe side effects. Using this heuristic, she might select a different treatment from that chosen by the first patient. A third patient does not search for any information about the treatment and simply asks his doctor for a recommendation. Tailored information needs to be designed to satisfy the consumer's information search and decision strategy. Research about individual differences is still rare but is crucial to understanding the interplay between risk communication and cognitive strategies.

Integrating Information Sources

Another research branch addresses the interaction of different information sources and its effect on health decisions. To learn about health treatments, a patient can consult many different sources: health professionals, friends, patient associations, newspaper reports, and the World Wide Web. Especially the latter provides a new and prominent platform to learn more about health treatments. The Web has both advantages and disadvantages. There are almost no limits on what information patients can search for, and information is often up to date. However, at the same time, patients have to evaluate this information and judge the credibility and trustworthiness of various sources, which is particularly important as even short exposure to misinforming Web sites can have a lasting influence on people's risk perception (Betsch et al. 2010). Patients must decide to what extent they can, for instance, trust information from the pharmaceutical industry or how objective the information is on vaccine-critical Web sites. Evidence and opinions on the Web are often conflicting and might confuse rather than educate the consumer. Again, patients run the risk of being misinformed and misled if they rely on doubtful evidence. Research can help us understand people's Web information search behavior (e.g., Hargittai et al. 2010) and how people identify reliable sources. Also, research has started to identify the relevant skills required to use the Internet successfully (Hargittai 2005, 2009), which could lead to the development of interventions to improve those skills in the future.

In addition, the Web has had a direct impact on the doctor–patient relationship (e.g., Diaz et al. 2002; McMullan 2006). Patients are not "naïve" and uninformed any more but have prior attitudes and expectations when meeting their doctors. On the one hand, the doctor–patient interaction benefits from informed patients. Patients and doctors have a more balanced relationship and need less time since the patient is already informed. On the other hand, doctors might have difficulties disabusing patients of potentially false beliefs and expectations acquired through the Internet. The pros and cons of using the Internet to inform patients and increase expertise in laypeople still needs further research.

Implementing Theories of Risk Communication

Another issue that researchers have to address that relates to the intersection between research and practice is the lack of theoretical frameworks in the field of risk communication, and in decision aids in particular. Durand et al. (2008) evaluated 50 decision support technologies with respect to their theoretical frameworks. Only 17 were explicitly based on a theoretical framework, mainly expected-utility theory. The lack of theories and their application in the design of decision aids, as well as in the evaluation process, leads to insufficient and ad hoc constructed decision aids. Therefore, researchers need (1) to formulate and test theories on which decision support technologies can be based, (2) to make these theories available to designers of decision support technologies, and (3) to evaluate the implementation of the theories in the decision support technologies.

Obstacles to Implementing Risk Communication

Investigating and identifying transparent risk communication formats is one requirement to improve risk communication in society. However, the second step is to transfer research into practice. As the lack of theories in designing decision support technologies shows, there is a long way to go from theory to practice.

Policy makers still communicate relative instead of absolute risk reductions, pharmaceutical industries promote their interests with misleading statistics, and health professionals themselves have difficulties with numbers. Why is this still the case? An answer is found in the seven sins identified by Gigerenzer and Gray (2011), which is already mentioned in the ➔ Introduction. Three of these sins directly address the issue of transparent risk communication, and were part of this chapter: biased reporting in medical journals, pamphlets, and the media. We also alluded to the lack of statistical literacy in health professionals and will discuss consequences and challenges in the next section. We have also already described another sin: defensive decision making. Doctors often do not prescribe those treatments that seem best but are guided by the desire to minimize potential legal consequences. Finally, biased funding refers to the pharmaceutical industry often sponsoring research trials. Consequently, researchers are not free in the research topics they select, the study design, data analysis, and data interpretation since industrial interests need to be considered. One might argue that researchers have to disclose conflicts of interests. However, this is not always the case. Weinfurt et al. (2008) found that consistent disclosure of financial interests was the exception in the biomedical literature and asked editors and authors to take responsibility.

One movement that has tried to eliminate industrial interests in research trials was launched by *JAMA*, the *Journal of the American Medical Association*. The editors launched a requirement for independent statistical analysis of industry-driven research in 2005. In contrast to other medical journals, *JAMA* published fewer RCTs and also fewer industry-driven RCTs (Wager et al. 2010). Further advances have been made by the introduction of standards and guidelines for reporting observational research (i.e., STROBE statement 2007; CONSORT statement 2009). Simple checklists help authors and editors evaluate research reports and assure complete and unbiased reporting. Such guidelines and standards should not be restricted to scientific journals but should also be set for media health coverage and health advertisements.

Media analyses and advertisement content analyses have repeatedly shown that the media rarely communicate numbers, and when they do, they use biased formats. Yet the mass media have the power to shape health decisions (Grilli et al. 2009) and thereby intentionally or unintentionally misinform and mislead the public. Health promotion campaigns from the pharmaceutical industry primarily follow financial interests and persuade rather than inform. For instance, in the United States in 2008 the pharmaceutical industry ranked second (behind the automotive industry) in dollars spent on advertising (Nielsen 2009).

An alternative approach to advertising is the use of so-called facts boxes (Schwartz et al. 2007). Facts boxes summarize the current state of evidence about drugs or other treatments in a way that laypeople can easily understand. They cover basic information and provide numbers in transparent formats by contrasting treatment and placebo groups and hence serve an educational purpose. *Figure 24.1* represents a facts box with basic information about mammography screening based on current scientific evidence.

Teaching Statistical Literacy

Many people in our society are statistically illiterate and innumerate. This phenomenon applies not only to laypeople, but also to experts. A way out of this dilemma is to promote education in statistical thinking on at least three levels: Statistics should be taught in schools, and statistics training should be offered to health professionals and (science) journalists.

Statistical Teaching in Schools

Statistical thinking is hardly taught at schools. Mathematics curricula do not include teaching statistical concepts; instead, the focus is on the mathematics of certainty, such as algebra, geometry, and trigonometry. In contrast to a widespread belief that children cannot deal with statistics, children at the elementary school level are already capable of understanding fundamental concepts of statistical thinking, such as natural frequencies and icon arrays (Zhu and Gigerenzer 2006). Hands-on approaches to problem solving, such as with tinker cubes, lego-like units, allow even first graders to learn about conditional frequencies through play (Kurz-Milcke and Martignon 2007; Kurz-Milcke et al. 2008). Despite attempts to include statistics in school curricula, there are four constraints that undermine successful and sustainable implementation. First, the first contact with statistics occurs too late in schools. Second, many textbooks use confusing representation formats. Third, statistics are often taught in a pallid way by abstract and unrealistic examples. Fourth, teachers themselves are often not as familiar with these concepts as they ought to be. A rethinking in mathematical teaching is pivotal for future statistically literate generations.

Statistics Training Education for Health Professionals

The second step addresses education of health professionals. Doctors directly interact with patients and therefore require the skills not only to understand statistics, but also to transparently communicate them. As far back as 1937 an editorial in the Lancet called attention to the strong link between medicine and statistics and the lack of fundamental abilities of doctors to deal with statistical information ("Mathematics and Medicine" 1937). It stated that the use (or abuse) of statistics "tends to induce a strong emotional reaction in non-mathematical minds." It complained that for "most of us figures impinge on an educational blind spot," which "is a misfortune, because simple statistical methods concern us far more closely than many of the things that we are forced to learn in the 6 long years of the medical curriculum." What has changed since then? Doctors still have trouble calculating positive predictive values and are prone to framing effects (e.g., 5-year survival rates vs. mortality rates, or relative vs. absolute risk reduction). Health professionals need to be trained in statistics. This will teach them how to identify biased reporting and how to translate statistical information into transparent formats.

Statistics Education for (Science) Journalists

The third target population is journalists. As previously mentioned, the mass media play an important role in educating the public. However, journalists might just reproduce biased reporting that has its origin in the medical literature. Therefore, educating scientists and making them aware of these biases will help them see through embellishments and obfuscations to translate risk information into comprehensive formats. They may also put public pressure on those who practice biased reporting.

Conclusion

Risk communication is a requirement for an informed public to be able to adequately deal with risks and uncertainties. On the one hand, experts and laypeople have difficulties in dealing with statistical information. On the other hand, the problem is less in people's minds and more in a health environment that puts little effort into presenting risks in an unbiased way. Biased reporting encompasses the omission of important information as well as the use of nontransparent communication formats. Informed and shared decision making will remain an illusion unless transparent risk communication formats are consistently applied.

We believe that statistically literate patients improve health decisions on an individual as well as on a public health level. Throughout this chapter, we proposed ways to design risk communication to educate and inform patients, instead of persuading them. These points should be kept in mind (see *Table 24.3*):

– Absolute risk changes are preferred over relative risk changes.
– Natural frequencies facilitate Bayesian reasoning in comparison to conditional probabilities.
– Annual mortality rates are less misleading and less biased than 5-year survival rates.
– Graphs can help overcome innumeracy.
– Disclosing uncertainty can help overcome the illusion of certainty.

People are able to make personal decisions that reflect their preferences and needs when they have sufficient information on which to base their decisions. There are two fundamental "adjustment screws": the consequent application of transparent communication formats and the implementation of education programs on different societal levels. Last but not least,

Table 24.3

Nontransparent versus transparent communication of risks: Four examples of how risks can be communicated to mislead and misinform the public and their transparent counterparts

How to communicate risks *nontransparently*	How to communicate risks *transparently*
Relative risks "The new generation of the contraceptive pill increases the risk of thrombosis by 100%."	*Absolute risks* "The new generation of the contraceptive pill increases the risk of thrombosis from 1 in 7,000 to 2 in 7,000."
Conditional probabilities – The probability of breast cancer is 1% for a woman at age 40 who participates in routine screening (this is the prevalence or base rate) – If a woman has breast cancer, the probability is 90% that she will get a positive mammography (this is the sensitivity or hit rate) – If a woman does not have breast cancer, the probability is 9% that she will also get a positive mammography (this is the false-positive rate) What is the probability that a woman at age 40 who had a positive mammogram actually has breast cancer? $P(H\|D) = \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.09 \cdot 0.99} = 0.092$	*Natural frequencies* – Ten out of 1,000 women at age 40 who participate in mammography screening have breast cancer (prevalence or base rate) – Of these 10 women, 9 have a positive mammogram (sensitivity or hit rate) – Out of the 990 women who do not have breast cancer, about 89 will have a positive mammogram nonetheless (false-positive rate) Now imagine a representative sample of 1,000 women age 40 who participate in breast cancer screening. How many of these women with a positive test result actually have breast cancer? $P(H\|D) = \frac{9}{9+89} = 9.2$
Five-year survival rate "The 5-year survival rate for people diagnosed with prostate cancer is 98% in the USA vs. 71% in Britain."	*Annual mortality rate* "There are 26 prostate cancer deaths per 100,000 American men versus 27 per 100,000 men in Britain."
Single-event probability "If you take Prozac, the probability that you will experience sexual problems is 30–50% (or: 30 to 50 chances out of 100)."	*Frequency statement* "Out of every 10 of my patients who take Prozac, 3–5 experience a sexual problem."

lessons learned in health risk communication can be adapted to other domains as well. Transparency and statistical literacy help people evaluate financial, environmental, and technological risks, and enable society to competently meet future challenges.

References

Allen M, Preiss R (1997) Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. Commun Res Rep 14:125–131

Ancker JS, Kaufman D (2007) Rethinking health numeracy: a multidisciplinary literature review. J Am Med Inform Assoc 14:713–721

Ancker JS, Senathirajah Y, Kukafka R, Starren JB (2006) Design features of graphs in health risk communication: a systematic review. J Am Med Inform Assoc 3:608–618

Baesler JE (1997) Persuasive effects of story and statistical evidence. Argument Advocacy 33:170–175

Baesler JE, Burgoon JK (1994) The temporal effects of story and statistical evidence on belief change. Commun Res 21:582–602

Barbey AK, Sloman SA (2007) Base-rate respect: from ecological rationality to dual processes. Behav Brain Sci 30:241–297

Barton A, Mousavi S, Stevens JR (2007) A statistical taxonomy and another "chance" for natural frequencies. Behav Brain Sci 30:255–256

Berry D, Raynor T, Knapp P, Bersellini E (2004) Over the counter medicines and the need for immediate action: a further evaluation of European commission recommended wordings for communicating risk. Patient Educ Couns 53:129–134

Betsch C, Renkewitz F, Betsch T, Ulshöfer C (2010) The influence of vaccine-critical Internet pages on perception of vaccination risks. J Health Psychol 15:446–455

Bodemer N, Müller SM, Okan Y, Garcia-Retamero R, Neumeyer-Gromen A (2011) Do the media provide transparent health information? A cross-cultural comparison of public information about the HPV vaccine (Submitted)

Brase GL (2009) Pictorial representations in statistical reasoning. Appl Cogn Psychol 23:369–381

Brun W, Teigen KH (1988) Verbal probabilities: ambiguous, context-dependent, or both? Organ Behav Hum Decis Process 41:390–404

Bucher HC, Weinbacher M, Gyr K (1994) Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. Br Med J 309:761–764

Budescu DV, Wallsten TS (1985) Consistency in interpretation of probabilistic phrases. Organ Behav Hum Decis Process 36:391–405

Casscells W, Schoenberger A, Grayboys T (1978) Interpretation by physicians of clinical laboratory results. N Engl J Med 299:999–1000

Consort (2009) Consolidated standards of reporting trials. www.consort-statement.org. Accessed April 2011

Cosmides L, Tooby J (1996) Are humans good intuitive statisticians after all? Rethinking some conclusions of the literature on judgment under uncertainty. Cognition 58:1–73

Covey J (2007) A meta-analysis of the effects of presenting treatment benefits in different formats. Med Decis Making 27:638–654

Davids SL, Schapira MM, McAuliffe TL, Nattinger AB (2004) Predictors of pessimistic breast cancer risk perception in a primary care population. J Gen Intern Med 19:310–315

deWit JBF, Das E, Vet R (2008) What works best: objective statistics or a personal testimonial? An assessment of the persuasive effects of different types of message evidence on risk perception. Health Psychol 27:110–115

Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW (2002) Patients' use of the Internet for medical information. J Gen Intern Med 17:180–185

Dieckmann NF, Slovic P, Peters E (2009) The use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. Risk Anal 29:1473–1488

Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B et al (2010) Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. Br Med J 341:c4543

Dören M, Gerhardus A, Gerlach FM, Hornberg C, Kochen MM, Kolip P et al. (2008) Wissenschaftler/innen fordern Neubewertung der HPV-Impfung und ein Ende der irreführenden Informationen. http://www.uni-bielefeld.de/gesundhw/ag3/downloads/Stellungnahme_Wirksamkeit_HPV-Impfung.pdf. Accessed April 2011

Durand M-A, Stiel M, Boivin J, Elwyn G (2008) Where is the theory? Evaluating the theoretical frameworks described in decision support technologies. Patient Educ Couns 71:125–135

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 246–267

Edwards A, Elwyn G (2009) Shared decision making in health care: achieving evidence-based patient choice. Oxford University Press, Oxford

Edwards A, Elwyn G, Covey J, Matthews E, Pill R (2001) Presenting risk information – A review of the effects of "framing" and other manipulations on patient outcomes. J Health Commun 6:61–82

Einhorn HJ, Hogarth RM (1985) Ambiguity and uncertainty in probabilistic inference. Psychol Rev 92:433–461

Epstein LG (1999) A definition of uncertainty aversion. Rev Econ Stud 66:579–608

Erev I, Cohen BL (1990) Verbal versus numerical probabilities: efficiency, biases, and the preference paradox. Organ Behav Hum Decis Process 45:1–18

Estrada CA, Martin-Hryniewicz M, Peek BT, Collins C, Byrd JC (2004) Literacy and numeracy skills and anticoagulation control. Am J Med Sci 328:88–93

European Commission (1998) A guideline on the readability of the label and package leaflet of medicinal products for human use. EC Pharmaceuticals Committee, Brussels

Fagerlin A, Wang C, Ubel PA (2005) Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? Med Decis Making 25:398–405

Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM (2007) Measuring numeracy without a math test: development of the subjective numeracy scale. Med Decis Making 27:672–680

Feufel M, Antes G, Gigerenzer G (2010) Competence in dealing with uncertainty lessons to learn from the influenza pandemic (H1N1) 2009. Bundesgesundheitsblatt 53:1283–1289

Frewer LJ (1999) Risk perception, social trust, and public participation in strategic decision making: implications for emerging technologies. Ambio 28:569–574

Frewer LJ, Hunt S, Brennan M, Kuznesof S, Ness M, Ritson C (2003) The views of scientific experts on how the public conceptualize uncertainty. J Risk Res 6:75–85

Frosch DL, Krueger PM, Hornik RC, Cronholm PF, Barg FK (2007) Creating demand for prescription drugs: a content analysis of television direct-to-consumer advertising. Ann Fam Med 5:6–13

Gaissmaier W, Straubinger N, Funder DC (2007) Ecologically structured information: the power of pictures and other effective data presentations. Behav Brain Sci 30:263–264

Gal I (1995) Big picture: what does "numeracy" mean? http://mathforum.org/teachers/adult.ed/articles/gal.html. Accessed April 2011

Galesic M, Garcia-Retamero R (2010) Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples. Arch Intern Med 170:462–468

Galesic M, Garcia-Retamero R (2011) Graph literacy: a cross-cultural comparison. Med Decis Making 31:444–457

Galesic M, Garcia-Retamero R, Gigerenzer G (2009) Using icon arrays to communicate medical risks: overcoming low numeracy. Health Psychol 28:210–216

Garcia-Retamero R, Galesic M (2009) Communicating treatment risk reduction to people with low numeracy skills: a cross-cultural comparison. Am J Public Health 99:2196–2202

Garcia-Retamero R, Galesic M (2010) Who profits from visual aids: overcoming challenges in people's understanding of risks. Soc Sci Med 70:1019–1025

Garcia-Retamero R, Galesic M, Gigerenzer G (2010) Do icon arrays help reduce denominator neglect? Med Decis Making 30:672–684

Gigerenzer G (2002) Calculated risks: how to know when numbers deceive you. Simon & Schuster, New York

Gigerenzer G (2007) Gut feeling: the intelligence of the unconscious. Viking, New York

Gigerenzer G, Gray M (2011) Launching the century of the patient. In: Gigerenzer G, Gray M (eds) Better doctors, better patients, better decisions: envisioning health care 2020. MIT Press, Cambridge, pp 3–28

Gigerenzer G, Hoffrage U (1995) How to improve Bayesian reasoning without instruction: frequency formats. Psychol Rev 102:684–704

Gigerenzer G, Hoffrage U, Ebert A (1998) AIDS counselling for low-risk clients. AIDS Care 10:197–211

Gigerenzer G, Hertwig R, van den Broek E, Fasolo B, Katsikopoulos KV (2005) "A 30% chance of rain tomorrow": how does the public understand probabilistic weather forecasts? Risk Anal 25:623–629

Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S (2007) Helping doctors and patients to make sense of health statistics. Psychol Sci Public Interest 8:53–96

Gigerenzer G, Mata J, Frank R (2009) Public knowledge of benefits of breast and prostate cancer screening in Europe. J Natl Cancer Inst 101:1216–1220

Gigerenzer G, Wegwarth O, Feufel M (2010) Misleading communication of risk: editors should enforce transparent reporting in abstracts. Br Med J 341:c4830

Golbeck AL, Ahlers-Schmidt CR, Paschal AM, Dismuke S (2005) A definition and operational framework for health numeracy. Am J Prev Med 29:375–376

Grigg W, Donahue P, Dion G (2007) The nation's report card:12th-grade reading and mathematics 2005 (NCES Report No. 2007–468). U.S. Department of Education, National Center for Education Statistics, Washington, DC

Grilli R, Ramsey C, Minozzi S (2009) Mass media interventions: effects on health care utilization. Cochrane Database Syst Rev 1:CD000389

Gurmankin AD, Baron J, Armstrong K (2004) The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. Med Decis Making 24:265–271

Hacking I (1975) The emergence of probability. Cambridge University Press, Cambridge

Hargittai E (2005) Survey measures of web-oriented digital literacy. Soc Sci Comput Rev 23:371–379

Hargittai E (2009) An update on survey measures of web-oriented digital literacy. Soc Sci Comput Rev 27:130–137

Hargittai E, Fullerton F, Menchen-Trevino E, Thomas K (2010) Trust online: young adults' evaluation of web content. Int J Commun 4:468–494

Heesen C, Köpke S, Kasper J, Richter T, Beier M, Mühlhauser I (2008) Immuntherapien der Multiplen Sklerose. http://www.zmnh.uni-hamburg.de/martin/dl/pinfo/immuntherapien_ms_inims_hamburg.pdf. Accessed April 2011

Heesen C, Kleiter I, Nguyen F, Schäffler N, Kasper J, Köpke S et al (2010) Risk perceptions in natalizumab-treated multiple sclerosis patients and their neurologists. Mult Scler 16:1507–1512

Hembroff LA, Holmes-Rovner M, Wills CE (2004) Treatment decision-making and the form of risk communication: results of a factorial survey. BMC Med Inform Decis Mak 4:1184–1120

Hoffrage U, Gigerenzer G, Krauss S, Martignon L (2002) Representation facilitates reasoning: what natural frequencies are and what they are not. Cognition 84:343–352

Holmes BJ, Henrich N, Hancock S, Lestou V (2009) Communicating with the public during health crises: experts' experiences and opinions. J Risk Res 12:793–807

Ibrekk H, Morgan MG (1987) Graphical communication of uncertain quantities to nontechnical people. Risk Anal 7:519–529

Johnson BB, Slovic P (1995) Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust. Risk Anal 15: 485–494

Kahneman D, Tversky A (1972) Subjective probability: a judgment of representativeness. Cognit Psychol 3:430–454

Kirsch IS, Jungeblut A, Jenkins L, Kolstad A (2007) Adult literacy in America: a first look at the findings of the National Adult Literacy Survey (NCES Report No. 1993–275; 3rd edn, U.S. Department of Education, National Center for Education Statistics. Washington, DC

Knapp P, Raynor DK, Berry DC (2004) Comparison of two methods of presenting risk information to patients about the side effects of medicines. Qual Saf Health Care 13:176–180

Knight FH (1921) Risk, uncertainty and profit. Harper, New York

Kotler P, Lee NR (2007) Social marketing: influencing behaviors for good. Sage, London

Kuhn KM (2000) Message format and audience values: interactive effects of uncertainty information and environmental attitudes on perceived risk. J Environ Psychol 20:41–51

Kurzenhäuser S (2003) Welche Informationen vermitteln deutsche Gesundheitsbroschüren über die Screening-Mammographie? Z Arztl Fortbild Qualitatssich 97:53–57

Kurz-Milcke E, Martignon L (2007) Stochastische Urnen und Modelle in der Grundschule (Stochastic urns and models in elementary school). In: Kaiser G (ed) Tagungsband der Jahrestagung der Gesellschaft für Didaktik der Mathematik. Verlag Franzbecker, Berlin

Kurz-Milcke E, Gigerenzer G, Martignon L (2008) Transparency in risk communication: graphical and analog tools. In: Tucker WT, Ferson S, Finkel A, Long TF, Slavin D, Wright P (eds) Strategies for risk communication: evolution, evidence, experience, vol 1128, Annals of the New York Academy of Sciences. Blackwell, New York, pp 18–28

Kutner M, Greenberg E, Jin Y, Paulsen C (2006) The health literacy of America's adults: results from the 2003 national assessment of adult literacy (NCES Report No. 2006–483). Government Printing Office, Washington, DC

Lafata JE, Simpkins J, Lamerato L, Poisson L, Divine G, Johnson CC (2004) The economic impact of false-positive cancer screens. Cancer Epidemiol Biomarkers Prev 13:2126–2132

Lewis C, Keren G (1999) On the difficulties underlying Bayesian reasoning: comment on Gigerenzer and Hoffrage. Psychol Rev 106:411–416

Lipkus IM (2007) Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. Med Decis Making 27:696–713

Lipkus IM, Hollands JG (1999) The visual communication of risks. J Natl Cancer Inst Monogr 25:149–162

Lipkus IM, Peters E (2009) Understanding the role of numeracy in health: proposed theoretical framework and practical insight. Health Educ Behav 36:1065–1081

Lipkus IM, Samsa G, Rimer BK (2001) General performance on a numeracy scale among highly educated samples. Med Decis Making 21:37–44

Macchi L, Mosconi G (1998) Computational features vs frequentist phrasing in the base-rate fallacy. Swiss J Psychol 57:79–85

Malenka DJ, Baron JA, Johansen S, Wahrenberger JW, Ross JM (1993) The framing effect of relative versus absolute risk. J Gen Intern Med 8:543–548

Marcus PM, Bergstrahl EJ, Zweig MH, Harris A, Offord KP, Fontana RS (2006) Extended lung cancer incidence follow-up in the Mayo lung project and overdiagnosis. J Natl Cancer Inst 98:748–756

Marteau TM, Saidi G, Goodburn S, Lawton J, Michie S, Bobrow M (2000) Numbers or words? A randomized controlled trial of presenting screen negative results to pregnant women. Prenat Diagn 20:714–718

Mathematics and medicine (1937, January 2). Lancet i:31

Mazur DJ, Hickham DH, Mazur MD (1999) How patients' preferences for risk communication influence treatment choice in a case of high risk and high therapeutic uncertainty: asymptotic localized prostate cancer. Med Decis Making 19: 394–398

McMullan M (2006) Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. Patient Educ Couns 63:24–28

Moxey A, O'Connell D, McGettigan P, Henry D (2003) Describing treatment effects to patients: how they are expressed makes a difference. J Gen Intern Med 18:948–959

Moynihan R, Bero L, Ross-Degnan D, Henry D, Lee K, Watkins J et al (2000) Coverage by the news media of the benefits and risks of medications. New Engl J Med 342:1645–1650

Mühlhauser I, Kasper J, Meyer G (2006) FEND: understanding of diabetes prevention studies: questionnaire survey of professionals in diabetes care. Diabetologia 49:1742–1746

Nadav-Greenberg L, Joslyn S (2009) Uncertainty forecasts improve decision making among nonexperts. J Cogn Eng Decis Making 3:209–227

Naylor CD, Chen E, Strauss B (1992) Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? Ann Intern Med 171:916–921

Neumeyer-Gromen A, Bodemer N, Müller SM, Gigerenzer G (in press) Ermöglichen Medienberichte und Informationsbroschüren zur Gebärmutterhalskrebsprävention informierte Entscheidungen? Eine Medienanalyse in Deutschland (Submitted)

Nielsen Company (2009) U.S. ad spending fell 2.6% in 2008, Nielson reports. Nielson Company, New York

Nisbett RE, Ross LD (1980) Human inference: strategies and shortcomings of social judgment. Prentice-Hall, Englewood Cliffs

Paulos JA (1988) Innumeracy: mathematical illiteracy and its consequences. Hill & Wang, New York

Pepper S, Prytulak LS (1974) Sometimes frequently means seldom: context effects in the interpretation of quantitative expressions. J Res Pers 8:95–101

Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S (2006) Numeracy and decision making. Psychol Sci Public Interest 17:407–413

Politi MC, Han PKJ, Col NF (2007) Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making 27:681–695

Politi MC, Clark MA, Ombao H, Dizon D, Elwyn G (2010) Communicating uncertainty can lead to less decision satisfaction: a necessary cost of involving patients in shared decision making? Health Expect 14:1–8

Reinard JC (1988) The empirical study of the persuasive effects of evidence: the status after fifty years of research. Hum Commun Res 15:3–59

Reyna V, Nelson WL, Han PK, Dieckmann NF (2009) How numeracy influences risk comprehension and medical decision making. Psychol Bull 135: 943–973

Rothman RL, Housam R, Weiss H, Davis D, Gregory R, Gebretsadik T, Shintani A, Elasy TA (2006) Patient understanding of food labels: the role of literacy and numeracy. Am J Prev Med 31:391–398

Sandblom G, Varenhorst E, Rosell J, Löfman O, Carlsson P (2011) Randomised prostate cancer screening trial: 20 year follow-up. Br Med J 342:d1539

Sarfati D, Howden-Chapman P, Woodward A, Salmond C (1998) Does the frame affect the picture? A study into how attitudes to screening for cancer are affected by the way benefits are expressed. J Med Screen 5:137–140

Schapira MM, Nattinger AB, McHorney CA (2001) Frequency or probability? A qualitative study of risk communication formats used in health care. Med Decis Making 21:459–467

Schröder FH, Hugosson J, Roobol MJ et al (2009) Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 360: 1320–1328

Schwartz LM, Woloshin S, Black WC, Welch HG (1997) The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 127:966–972

Schwartz LM, Woloshin S, Dvorin EL, Welch HG (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, Welch HG (2007) Using a drug facts box to communicate drug benefits and harms. Ann Intern Med 150:516–527

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

Shaw NJ, Dear PRF (1990) How do parents of babies interpret qualitative expressions of probability. Arch Dis Child 65:520–523

Sheridan S, Pignone MP, Lewis CL (2003) A randomized comparison of patients' understanding of number needed to treat and other common risk reduction formats. J Gen Intern Med 18:884–892

Smeeth L, Haines A, Ebrahim S (1999) Numbers needed to treat derived from meta-analyses – sometimes informative, usually misleading. Br Med J 318:1548–1551

Steckelberg A, Balgenorth A, Mühlhauser I (2001) Analyse von deutschsprachigen Verbraucher-Informationsbroschüren zum Screening auf kolorektales Karzinom. Z Arztl Fortbild Qualitatssich 95:535–538

Steckelberg A, Berger B, Köpke S, Heesen C, Mühlhauser I (2005) Criteria for evidence-based patient information. Z Arztl Fortbild Qualitatssich 99:343–351

Steurer J, Held U, Schmidt M, Gigerenzer G, Tag B, Bachmann L (2009) Legal concerns trigger prostate-specific antigen testing. J Eval Clin Pract 15:390–392

Stone ER, Yates JF, Parker AM (1997) Effects of numerical and graphical displays on professed risk-taking behavior. J Exp Psychol Appl 3:243–256

Stone ER, Sieck WR, Bull BE, Yates JF, Parks SC, Rush CJ (2003) Foreground:background salience: explaining the effects of graphical displays on risk avoidance. Organ Behav Hum Decis Process 90:19–36

Strobe-Statement (2007) Strengthening the reporting of observational studies in epidemiology. http://www.strobe-statement.org/. Accessed April 2011

Taylor SE, Thompson SC (1982) Stalking the elusive "vividness" effect. Psychol Rev 89:155–181

Thompson KM (2002) Variability and uncertainty meet risk management and risk communication. Risk Anal 22:647–654

Ubel PA, Jepson C, Baron J (2001) The inclusion of patient testimonials in decision aids: effects on treatment choices. Med Decis Making 21:60–68

van Dijk H, Houghton J, van Kleef E, van der Lans I, Rowe G, Frewer LJ (2008) Consumer responses to communication about food risk management. Appetite 50:340–352

Viscusi WK, Magat WA, Huber J (1991) Communication of ambiguous risk information. Theory Decis 31:159–173

Wager E, Mhaskar R, Warburton S, Djulbegovic B (2010) JAMA published fewer industry-funded studies after introducing a requirement for independent statistical analysis. PLoS One 5:e13591

Wainer H (1984) How to display data badly. Am Stat 38:137–147

Wallsten TS, Fillenbaum S, Cox JA (1986) Base rate effects on the interpretation of probability and frequency expressions. J Mem Lang 25:571–587

Wallsten TS, Budescu DV, Zwick R, Kemp SM (1993) Preferences and reasons for communicating probabilistic information in verbal or numerical terms. Bull Psychon Soc 31:135–138

Weber EU, Hilton DJ (1990) Contextual effects in the interpretations of probability words: perceived base rate and severity of events. J Exp Psychol Hum Percept Perform 16:781–789

Wegwarth O, Gaissmaier W, Gigerenzer G (2011) Deceiving and informing: the risky business of risk perception. Med Decis Making 31:378–379

Weinfurt KP, Seils DM, Tzeng JP, Lin L, Schulman KA, Califf RM (2008) Consistency of financial interest disclosures in the biomedical literature: the case of coronary stents. PLoS One 3:e2128

Welch HG, Schwartz LM, Woloshin S (2000) Are increasing 5-year survival rates evidence of success against cancer? J Am Med Assoc 283:2975–2978

Wilde J (2009) PSA screening cuts deaths by 20%, says world's largest prostate cancer study. ERSPC press office, carver wilde communications. http://www.erspc-media.org/release090318.php. Accessed April 2011

Witte K, Allen M (2000) A meta-analysis of fear appeals: implications for effective public health campaigns. Health Educ Behav 27:591–615

Woloshin S, Schwartz LM, Black WC, Welch HG (1999) Women's perceptions of breast cancer risk: how you ask matters. Med Decis Making 19:221–229

Zhu L, Gigerenzer G (2006) Children can solve Bayesian problems: the role of representation in mental computation. Cognition 98:287–308

Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A (2007) Validation of the subjective numeracy scale (SNS): effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making 27:663–671

Zikmund-Fisher BJ, Fagerlin A, Ubel PAU (2008a) Improving understanding of adjuvant therapy options by using simpler risk graphics. Cancer 113:3382–3390

Zikmund-Fisher BJ, Ubel PA, Smith DM, Derry HA, McClure JB, Stark AT et al (2008b) Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. Patient Educ Couns 73:209–214

Zimmer AC (1983) Verbal vs. numerical processing of subjective probabilities. In: Scholz RW (ed) Decision making under uncertainty. Elsevier, Amsterdam, pp 159–182