When Misinformed Patients Try to Make Informed Health Decisions

Wolfgang Gaissmaier and Gerd Gigerenzer

Abstract

Statistical illiteracy in health—the inability to understand health statistics—is widespread among the general public. Many people find it hard to accept uncertainty in the first place and, even if they do, basic numerical information is difficult to understand. The problem is aggravated when the benefits and harms of treatment options must be evaluated or test outcomes understood.

Statistical illiteracy results not only from a lack of education but from the nontransparent framing of information that is sometimes unintentional, but which can also be used deliberately to manipulate people. In health care, nontransparent framing of information seems to be the rule rather than the exception. Patients have difficulties finding reliable information—on the Internet, in invitations to screening, medical pamphlets, or media reports—yet this situation can be corrected. Statistical thinking must be taught to the public, and health care workers and journalists must be trained in transparent framing. Knowing what questions to ask, what information is missing, and how to translate nontransparent statistics into transparent ones would empower an educated citizenry to reject attempts to persuade rather than inform.

Introduction

A recent press release advertised the results of an international randomized controlled trial on the benefits of using PSA tests for prostate cancer screening. It stated that PSA screening would reduce the risk of dying from prostate cancer by 20% (Wilde 2009). But what does that really mean? Just how big is 20%, and how many people does it equate to? It meant that out of every 1,410 men who regularly participated in prostate cancer screening, 1 less person died of prostate cancer than in an equally large group of men who did not participate (Schröder et al. 2009). What the press release did not report, however, is that
out of these 1,410 men, 48 were unnecessarily treated and hence subjected to potential incontinence and impotence. In addition, it omitted the fact that there is no evidence that PSA screening reduces overall mortality at all. That is, the 1 man (out of 1,410) who was saved from dying from prostate cancer likely died from something else in the same time frame. The press release also failed to mention that there was a second large clinical trial, published in the same issue of that medical journal, where no reduction in prostate cancer mortality was found to be attributable to screening (Andriole et al. 2009).

This example illustrates two key problems in the communication of clinical evidence: communication is often made nontransparent through the use of, what is called, relative risks (“20%”) instead of more transparent absolute risks (“1 in 1,410”), and it is often incomplete as only those facts which suit the interests of the communicator are reported. Together with lack of statistical education, nontransparency and incomplete reporting contribute to what we call statistical illiteracy in health; that is, the inability to understand health statistics. One major cause of statistical illiteracy is conflict of interest on the part of those communicating health statistics, who push an agenda instead of informing citizens.

The problem of statistical illiteracy is a collective one (Gigerenzer 2002; Gigerenzer et al. 2007; Reyna and Brainerd 2007). It exists not only in patients but also in physicians, politicians, and journalists. In this chapter, we focus on the extent of statistical illiteracy in the general public, explain the major causes, and offer some remedies on how it could be overcome. We do not intend to provide a comprehensive survey, but use evidence to exemplify the general situation in health care.

**Extent of Statistical Literacy among Patients**

To be health literate, patients need to have at least four basic abilities. They need to (a) accept living with uncertainty, (b) have a basic understanding of numerical information, (c) grasp the benefits and harms of treatment options, and (d) understand test results.

**The Illusion of Certainty**

To appreciate the importance of health statistics, patients need to understand that in the first place, there is no such thing as certainty. As Benjamin Franklin once said: “In this world, there is nothing certain but death and taxes.” Yet living with uncertainty can be extremely difficult when health statistics, as opposed to baseball statistics, are at issue. The term “illusion of certainty” refers to an emotional need for certainty when none exists. This feeling can be attached to test results that are taken to be absolutely certain and to treatments that appear to guarantee a cure. One might think that physicians routinely
inform patients that even the best tests are not perfect and that every test result needs therefore to be interpreted with care, or may need to be repeated. Studies indicate, however, that clinicians rarely communicate the uncertainties about risks and benefits of treatments to patients (Braddock et al. 1999).

In a nationwide survey conducted in 2006, 1,000 German citizens over the age of 18 were asked: “Which of the following tests are absolutely certain?” While only 4% believed that an expert horoscope could give absolutely accurate results, a majority of Germans believed that HIV tests, fingerprints, and DNA tests were absolutely certain, even though none of these are (Gigerenzer et al. 2007). In contrast to these tests, which tend to make relatively few errors, a mammography (positive or negative mammogram) has a miss rate of about 10%, and its false positive rate is nearly as high. Nonetheless, it was rated as “absolutely certain” by 46% of the women and 42% of the men. Higher education is only a slight safeguard against the illusion of certainty: One-third of women with a university degree also believed that mammograms are absolutely certain.

Once one accepts that uncertainty is unavoidable in the world, the next challenge is to understand numbers that express uncertainties. This most elementary skill is called basic numeracy.

**Basic Numeracy**

Schwartz, Woloshin, Black, and Welch (1997) developed a simple three-question scale to measure basic numeracy:

1. A person taking Drug A has a 1% chance of having an allergic reaction. If 1,000 people take Drug A, how many would you expect to have an allergic reaction? ____ person(s) out of 1,000
2. A person taking Drug B has a 1 in 1,000 chance of an allergic reaction. What percent of people taking Drug B will have an allergic reaction? ____ %
3. Imagine that I flip a coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips? ____ times out of 1,000

The test was applied to a random sample of female veterans in New England, 96% of whom were high school graduates and whose average age was 68. Only 54% were able to convert 1% to 10 in 1,000; 20% were able to convert 1 in 1,000 to 0.1%; and 54% were able to answer that one would expect the coin to come up heads 500 times out of 1,000, with the most common incorrect answers to the last being 25, 50, and 250. The number of correct answers was strongly related to women’s ability to accurately interpret the benefit of mammography after being presented with standard risk reduction information: Only 6% of women answering just one basic numeracy question correctly could accurately interpret the data, compared to 40% of those answering all three
W. Gaissmaier and G. Gigerenzer

questions correctly. Thus, basic numeracy seems to be a necessary precondition for making informed decisions (cf. Reyna et al. 2009), and although physicians do better on this test, even they are far from perfect (see Wegwarth and Gigerenzer, this volume).

Table 2.1 shows the prevalence of low numeracy skills in representative samples of the general population of U.S. and German adults. Note again the great difficulty that large parts of the public, just as the female veterans, had in translating small frequencies into percentages. The left column in Table 2.1 shows that only 25% of U.S. citizens correctly converted 1 in 1,000 to 0.1%. Even among the highest education groups, this percentage was merely 27%. In another study, Lipkus, Samsa, and Rimer (2001) found that only 21% of well-educated adults were able to calculate this problem correctly. Comparable results have been found recently in a survey of probabilistic, representative national samples in the United States and in Germany (Table 2.1, columns 3 and 4), and this question was again found to be the most difficult one. It was answered correctly by only 24% of U.S. and 46% of German adults. In addition, overall numeracy on an extended version of the original questionnaire was higher in Germany than in the United States. Education predicted numeracy in both countries, but the gap between those with low education (less than high school) and those with high education (college or more) was much larger in the United States (40% overall accuracy vs. 83%) than in Germany (62% vs. 81%).

Statistical literacy becomes even more important when interpreting the benefits and harms of treatments, which are often communicated using more complicated health statistics, such as relative risks. Recall that in the introductory example above, a 20% relative risk reduction (RRR) in prostate cancer

<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>Basic numeracy in the United States and Germany based on nationally representative samples. The table shows data from two studies and separately includes U.S. citizens with highest educational attainment for one study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. adults ages 35–70</td>
<td>Postgraduate degree</td>
</tr>
<tr>
<td>35–70</td>
<td>n = 450</td>
</tr>
<tr>
<td>Convert 1% → 10 in 1,000</td>
<td>70</td>
</tr>
<tr>
<td>Convert 1 in 1,000 → 0.1%</td>
<td>25</td>
</tr>
<tr>
<td>Heads in 1,000 coin flips</td>
<td>76</td>
</tr>
</tbody>
</table>

a Schwartz and Woloshin (2000, unpublished data)
b Galesic and Garcia-Retamero (2010)
mortality was the same as an absolute risk reduction (ARR) of 1 in 1,410. Are patients aware of the difference between these two formats?

**Understanding Benefits and Harms of Treatments**

Patients need realistic estimates of the benefits and harms of treatment options to be able to evaluate whether the benefits outweigh the harms. Often they do not have realistic estimates, however, as highlighted by a recent representative survey of more than 10,000 citizens from nine European countries (Gigerenzer et al. 2009). In this survey, using face-to-face interviews, men and women were asked to estimate the benefits of prostate and breast cancer screening, respectively. As shown in Figure 2.1, about 90% of both men and women overestimated the benefits of screening, by tenfold, hundredfold, or more, or did not know.

One reason for such misunderstandings is that benefits and harms of treatments are often communicated as relative risks, which, as we will illustrate, are difficult for people to understand. But there is a simple solution: Use absolute risks instead.

Is perceived treatment efficacy influenced by framing information in terms of RRR and ARR? In a telephone survey in New Zealand, respondents were given information on three different screening tests for unspecified cancers (Sarfati et al. 1998). The benefits were identical, except that they were expressed either as RRR, ARR, or the number of people needed to be treated (NNT) or screened to prevent 1 death from cancer (which is 1/absolute risk reduction):

- **RRR**: If you have this test every two years, it will reduce your chance of dying from this cancer by around one-third over the next ten years.
- **ARR**: If you have this test every two years, it will reduce your chance of dying from this cancer from around 3 in 1,000 to around 2 in 1,000 over the next 10 years.
- **NNT**: If around 1,000 people have this test every two years, 1 person will be saved from dying from this cancer every 10 years.

When the benefit of the test was presented as an RRR, 80% of 306 New Zealanders said they would likely accept the test. When the same information was presented as an ARR and NNT, only 53% and 43% wanted the test, respectively. Medical students also fall prey to the influence of framing (Naylor et al. 1992), as do patients (Malenka et al. 1993) and health professionals (Mühlhauser et al. 2006). In a meta-analysis on the effect of presenting information in terms of absolute risks versus relative risks, Covey (2007) analyzed 31 experiments that either investigated physicians, other health professionals, students, or patients which showed that all of them can be consistently manipulated by framing the treatment effect differently.
The problem with relative risks is that, unlike absolute risks, they remain silent about the baseline risk. A 50% risk reduction, for instance, can mean a reduction from 2 to 1 out of 100,000 people. However, it can also mean a reduction from 20,000 to 10,000 out of 100,000 people, which would be a

Figure 2.1 Results of a survey (Gigerenzer et al. 2009) on the perceived benefits of cancer screening in nine European countries (Austria, France, Germany, Italy, Netherlands, Poland, Russia, Spain, and the United Kingdom). Over 90% of men and women overestimated the benefits of mammography or PSA screening, or did not know. The data shows the results for the age group for which the screening is usually offered. (a) Perceived benefit of mammography screening by women age 50–69, summed across nine countries (n = 1467). Bars show the percentage of women who believed that out of 1,000 women participating in screening, X fewer will die from breast cancer compared to women who do not participate. The available evidence indicates that the best answer is 1 in 1,000 (Woloshin and Schwartz 2009). (b) Perceived benefit of PSA screening by men age 50–69, summed across the eight EU countries (n = 1291). Russia is excluded here, because Russian men estimated the benefits much more accurately than anyone else and thus would distort the EU picture here (for Russia, see Gigerenzer et al. 2009). Bars show the percentage of these men who believed that out of 1,000 men participating in screening, X fewer will die from prostate cancer compared to men who do not participate. Because estimates vary between 0 in 1,000 (Andriole et al. 2009) and 0.7 in 1,000 (= 1 in 1,410) (Schröder et al. 2009), we marked both 0 and 1 in 1,000 as correct.

The problem with relative risks is that, unlike absolute risks, they remain silent about the baseline risk. A 50% risk reduction, for instance, can mean a reduction from 2 to 1 out of 100,000 people. However, it can also mean a reduction from 20,000 to 10,000 out of 100,000 people, which would be a
Misinformed Patients Try to Make Informed Health Decisions

much more relevant effect. Particularly for low probability risks, communicating changes in relative terms makes effects loom larger than they actually are.

Understanding Test Results

Patients in a clinic in Colorado and Oklahoma were asked about standard tests for diseases such as strep throat infection, HIV, and acute myocardial infarction (Hamm and Smith 1998). Each patient judged (a) the probability that a person has the disease before being tested (prevalence or base rate), (b) the probability that a person tests positive if the disease is present (sensitivity), (c) the probability that a person tests negative if the disease is absent (specificity), and (d) the probability that a person has the disease if test results are positive (positive predictive value). Most patients estimated the four probabilities to be essentially the same—indeed, of whether the base rate was high or low, or the test accurate or not. Even experienced patients did not understand health statistics, which suggests that their doctors either never explained risks or failed to communicate them properly. Studies with university students show that they too have difficulties drawing conclusions from sensitivities and specificities (Gigerenzer and Hoffrage 1995).

The problem is that the relevant pieces of information, such as the chance of detecting a disease, are usually communicated as conditional probabilities (e.g., sensitivities and specificities). For instance, a 30-year-old pregnant woman took a nuchal scan for testing whether her unborn child has Down syndrome. She is told that the chance of a positive test result if the child actually has Down syndrome is 80% (sensitivity). That is, 80% is the conditional probability of testing positive given Down syndrome. However, this probability is often confused with the positive predictive value of the test; that is, the probability of the child having Down syndrome given a positive test result, which is, of course, not the same. This can be illustrated with a more intuitive example. Until now, every American president has been male; that is, the probability of being male given that one is president of the United States has been 100%. The reverse, obviously, does not hold: Given that one is male, chances of being the president of the United States are rather low.

The question is how to get from the sensitivity of the test to the positive predictive value, which is the information one really wants. Two further pieces of information are necessary. First, one needs to know the base rate of the disease; for a pregnant woman at the age of 30, this rate is about 0.15%. Second, one needs to know the false positive rate of the test; that is, the probability of getting a positive test result given that the child actually does not have Down syndrome, which is about 8%. Formally, the sensitivity, the base rate, and the false positive rate can be combined to calculate the positive predictive value by applying Bayes’s rule. However, both physicians and laypeople often have trouble with these probabilities, and it is much simpler to think about such problems in terms of what we call natural frequencies.
Instead of combining conditional probabilities such as sensitivities and false positive rates, natural frequencies are about concrete cases; for instance, 10,000 pregnant women. Out of these 10,000 women, we expect 15 to carry a child with Down syndrome (= 0.15% base rate); the remaining 9,985 women are not affected. Of the 15 women whose child has Down syndrome, we expect that 12 will receive a positive test result (= 80% sensitivity). From the remaining 9,985 women, about 799 women will also test (falsely) positive (= 8% false positive rate). Thus, there are 811 women who receive a positive test result, out of which 12 actually carry a child with Down syndrome. Therefore, the probability of a child having Down syndrome given a positive test—the positive predictive value—is about 12 out of 811, which is about 1.5% (Figure 2.2).

Estimating the probability of disease given a positive test (or any other posterior probability) has been repeatedly shown to be much easier with natural frequencies than with conditional probabilities in students (e.g., Gigerenzer and Hoffrage 1995, 1999), physicians (Gigerenzer et al. 2007), and the elderly (Galesic, Gigerenzer et al. 2009). Even fifth graders can consistently solve Bayesian problems with natural frequencies (Zhu and Gigerenzer 2006).

Note that natural frequencies refer to situations where two variables are considered: Natural frequencies are joint frequencies, as shown in Figure 2.2. In contrast, relative frequencies (which are numerically identical to the conditional probabilities in Figure 2.2) do not facilitate judgments (Gigerenzer and Hoffrage 1995), a fact sometimes misunderstood in the literature (Hoffrage et al. 2002).

<table>
<thead>
<tr>
<th>1 pregnant woman</th>
<th>10,000 pregnant women</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Down syndrome</strong></td>
<td><strong>no Down syndrome</strong></td>
</tr>
<tr>
<td>0.15%</td>
<td>99.85%</td>
</tr>
<tr>
<td>80% positive</td>
<td>8% positive</td>
</tr>
<tr>
<td>20% negative</td>
<td>92% negative</td>
</tr>
<tr>
<td>15</td>
<td>9,985</td>
</tr>
<tr>
<td>3</td>
<td>799</td>
</tr>
<tr>
<td>12</td>
<td>9,186</td>
</tr>
</tbody>
</table>

\[
p(\text{Down syndrome} | \text{test positive}) = \frac{0.0015 \times 0.8}{0.0015 \times 0.8 + 0.9985 \times 0.08} = \frac{12}{12 + 799}\]

**Figure 2.2** A 30-year-old pregnant woman gets a positive nuchal scan for Down syndrome. What are the chances that her child actually has Down syndrome (Kurzenhäuser and Hoffrage 2002)? The information in (a) is expressed in terms of conditional probabilities, and the formula is known as Bayes’s rule. The information in (b) is expressed in terms of natural frequencies, which simplify the computations and foster insight.
In summary, relative risks and conditional probabilities tend to confuse patients, whereas absolute risks (and NNTs) help them understand the benefits and harms, just as natural frequencies help communicate what a test result means.

**Causes**

There are various players in public health whose goals can conflict with transparent risk communication: pushing a political agenda, attracting media attention, selling a new drug, increasing compliance with screening, or trying to impress physicians. We will argue that a major reason for statistical illiteracy can thus be found not only in patients’ minds and lack of training, but in their environment. Statistical illiteracy is fostered by nontransparent framing of information that is sometimes an unintentional result of lack of understanding, but can also be an intentional effort to manipulate or persuade people.

As the above examples have shown, two number formats are particularly hard to understand for the general public: relative risks and conditional probabilities. These two formats can easily be replaced by transparent formats (absolute risks and natural frequencies), but unfortunately seldom are. Despite the general belief that one can find everything on the Internet or in other media, it is in fact difficult to find information about health care that is both transparent and complete. As Gigerenzer and Gray (this volume) emphasize, the problem already starts with nontransparent information in medical journals and is amplified by leaflets containing confusing, incomplete, or even wrong information. We will demonstrate that patients have a hard time finding good information, be it in searching through invitations to screening, medical pamphlets, the Internet, or media reports about medical research. All of these sources suffer from nontransparent information and are more likely to confuse people than to enlighten them.

**Are Patients Likely to Find Transparent Information in Medical Pamphlets or on the Internet?**

Patients only rarely consult first-hand sources such as medical journals. Instead, besides asking their physician, they rely on other sources of information. To illustrate the quality of information available to patients, consider the case of breast cancer screening. First of all, women who contemplate participating need some basic information about breast cancer. The mortality rate, not the incidence rate, is most relevant for screening, given that the goal of screening is to reduce mortality, whereas it cannot reduce the risk of getting breast cancer (incidence). Naturally, the incidence rates loom larger than the mortality rates and thus contribute to raising anxiety; campaigns that selectively report incidence rates have been criticized for this reason (Baines 1992). Most importantly, information on breast cancer screening should provide information
about the potential benefits and harms, so that a woman can make an informed
decision as to whether she wants to participate or not. Table 2.2 provides the
essential information about the benefits and harms. Benefits can be expressed
in terms of both total cancer mortality and breast cancer mortality. For women
who participate in screening, there is no difference in total cancer mortality,
but a reduction of breast cancer mortality of 1 in 1,000. At the same time,
screening may lead to two major harms: false positives resulting in unneces-
sary biopsies and unnecessary treatments, such as lumpectomies. A third harm
is getting radiation-induced cancer from the mammography, but since no reli-
able estimates of the actual danger exist, we have decided not to include this in
the facts box (Table 2.2).

The facts box is a transparent way to communicate the major benefits and
harms of screening and treatments in general (see Schwartz and Woloshin, this
volume). What information do women actually get on breast cancer screening,
and in what format do they get it? There are three sources of information we
focus on here: invitations to screening, pamphlets, and web sites. In countries
with publicly funded screening, eligible women are often made aware of these
programs by letters of invitation. Thus, by sheer numbers of people reached,
such letters are—alongside physicians—potentially the most important source
of information about screening. Invitation letters would be the ideal opportuni-
ty to provide patients with balanced, transparent information about screening,
so that they can make informed decisions. Jorgensen and Gøtzsche (2006) in-
vestigated such invitations to attend breast cancer screening in seven countries
with publicly funded screening: Australia, Canada, Denmark, New Zealand,
Norway, Sweden, and the United Kingdom (Table 2.3). Other studies have in-
vestigated the information that medical pamphlets and web sites provide about
breast cancer screening in many countries (Table 2.3).

Table 2.2 Facts box for mammography screening; potential benefits and harms are 10-

<table>
<thead>
<tr>
<th>Breast Cancer Screening with Mammography</th>
<th>Women aged 50 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1000 women not screened</td>
</tr>
<tr>
<td>Benefits?</td>
<td></td>
</tr>
<tr>
<td>• Total cancer mortality</td>
<td>No difference</td>
</tr>
<tr>
<td>• Breast cancer mortality</td>
<td>5</td>
</tr>
<tr>
<td>Harms?</td>
<td>4</td>
</tr>
<tr>
<td>• False positives with biopsies</td>
<td>—</td>
</tr>
<tr>
<td>• Overtreatmenta</td>
<td>—</td>
</tr>
<tr>
<td>Positive predictive value?b</td>
<td>1 out of 10</td>
</tr>
</tbody>
</table>

a For example, unnecessary lumpectomy.
b That is, proportion of women with breast cancer among those who test positive.
Most of the invitations (97%) stated the major benefit of screening is the reduction in breast cancer mortality. However, the very few (23%) that also mentioned the size of the benefit always did so by using RRRs rather than ARRs. None of the invitations included information about potential harms or the positive predictive value. Instead, most invitations used persuasive wording and prespecified appointments. Thus, the invitation letters clearly aim at compliance rather than at informing the public, which is probably a result of a built-in conflict: those who are responsible for the screening program are also responsible for designing the invitations, which puts their goal of increasing compliance at odds with increasing transparency.
In medical pamphlets, numbers are very often not given at all. In the few cases where quantitative estimates of the benefits are provided, they are almost exclusively presented as RRRs. Many of these sources of information also remain mute on the harms. However, some pamphlets (German and Austrian) do include information about harms from X-rays, even if mostly to reassure patients that these are negligible (Rásky and Groth 2004). Almost no pamphlet explains what a positive test result means.

Given this imbalance in reporting, it is probably not surprising that the survey conducted in nine European countries about the perceived benefits of cancer screening (Gigerenzer et al. 2009) found that those who said they sometimes or frequently consulted health information in pamphlets had no better understanding than those who did not. In Germany, where pamphlets were consulted by a record 41% of the public, those who consulted pamphlets sometimes or frequently showed even a higher overestimation of benefits than those who did not.

Web sites—in particular, those of advocacy groups and governmental institutions (24 web sites in total)—recommended screening and favored information that shed positive light on it. Only few mentioned the major potential harms of screening: overdiagnosis and overtreatment. Three web sites of consumer organizations had a more balanced perspective on breast cancer screening and included information on both the potential benefits and harms. In total, very few sites met the standards of informed consent, as specified by the General Medical Council’s (1998) guidelines for patient information. This problem is not limited to information about cancer and has been shown for other conditions as well (Impiccatore et al. 1997; Rigby et al. 2001). These results are alarming, given that many people use the Internet to acquire information about health issues—in the European Union, this number is 23% (see Jørgensen and Gøtzsche 2004).

Do the Media Provide Balanced Information about Health Topics?

The press has a powerful influence on public perceptions of health and health care; much of what people—including many physicians—know and believe about medicine comes from the print and broadcast media. However, journalism schools tend to teach everything except understanding numbers. Journalists generally receive no training in how to interpret or present medical research (Kees 2002). A survey of health reporters at daily newspapers in five U.S. Midwestern states (Voss 2002) reported that very few found it easy to interpret statistical data and less than a third found it easy to put health news in context. This finding is similar to that of the Freedom Forum survey, where nearly half of the science writers agreed that “reporters have no idea how to interpret scientific results” (Hartz and Chappell 1997:30).

Lack of education inevitably affects the quality of press coverage, particularly in the reporting of health statistics about medical research, as has been
repeatedly documented (e.g., Moynihan et al. 2000; Ransohoff and Harris 1997; Schwartz et al. 1999a). One disturbing problem associated with how the media report on new medications is the failure to provide quantitative data on how well the medication works. Instead, many of these news stories present anecdotes, often in the form of patients describing miraculous responses to the new medication. If the benefits of a medication are quantified at all, they are commonly reported as an RRR. The situation is similar when it comes to the harms of medications: Typically, less than half of stories name a specific side effect and even fewer actually quantify it. This is most dramatic in direct-to-consumer advertisements, which often display the RRR from the medication in prominent, large letters, but present harms in long lists in very fine print. Television ads typically give consumers more time to absorb information about benefits than about risks, resulting in better recall of benefits (Kaphingst et al. 2004, 2005).

In general, journalists do not seem to help the public understand health statistics, at least with regard to cancer screening. In none of the nine European countries surveyed by Gigerenzer et al. (2009), about the perceived benefits of cancer screening, were estimates of the benefits any better by those participants who reported sometimes or frequently using television or newspapers as a source for health information.

**Remedies**

The facts reviewed so far might be discouraging. It seems that many players in the field of health avoid transparency and complete reporting—ranging from the major medical journals (Nuovo et al. 2002; Schwartz et al. 2006) to governments and health organizations to the media. Yet this does not mean that there is nothing to be done. The most important means of improvement is teaching the public statistical thinking combined with training health care providers and journalists in transparent framing. But that necessitates rethinking how statistical thinking is taught.

**Teaching Statistical Literacy**

Statistical thinking is the most useful part of mathematics for life after school. In most countries, however, almost all of the available classroom time is spent on the mathematics of certainty—from algebra to geometry to trigonometry. If children learned to deal with an uncertain world in a playful way, much of collective statistical illiteracy would be history. The U.S. National Council of Teachers of Mathematics (NCTM) has announced its commitment to teaching data analysis and probability in grades prekindergarten to 12, as described in its Principles and Standards for School Mathematics (NCTM 2000), and declared data analysis and probability its “professional development focus of the
year,” providing additional resources and continuing education. The NCTM prefaced its Principles with a simple truth: “Young children will not develop statistical reasoning if it is not included in the curriculum.”

Today, the mathematics curriculum in many countries includes probability and statistics. Yet research on the effect of teaching showed that although students learn how to compute formal measures of averages and variability, even those with good grades rarely understand what these statistics represent or their importance and connection to other concepts (Garfield and Ben-Zvi 2007). Few learn to see a connection between statistics in school and what is going on in their world, which is exactly how one could grab their attention: by teaching real-world problems instead of throwing dice and flipping coins. Adolescents love baseball statistics, want to know about HIV, have heard about probabilities of rain, worry about the chance of a major earthquake, and when they get older, become interested in graphs about stock indices and develop concerns about cholesterol and blood pressure. How safe is the contraceptive pill? What is the error margin for polls and surveys? Personal relevance is what makes statistics so interesting. This would, of course, also require educating teachers first.

Teaching statistical literacy in school would lay the ground for also improving training doctors in medical statistics. Even back in 1937, an editorial in The Lancet noted 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 six long years of the medical curriculum.” Improvements have been made since then, although there are still medical organizations, physicians, and students who tend to see statistics as inherently mathematical and clinically irrelevant for the individual patient (Gigerenzer 2002; Altman and Bland 1991; Wegwarth and Gigerenzer, this volume).

Provide Complete and Transparent Information

An important response to statistical illiteracy is to give the public numbers. Patients have a right to learn how large the benefits and harms of a treatment are. Verbal descriptions of risks, however, are notoriously unclear and lead to misunderstandings (Steckelberg et al. 2005; Trevena et al. 2006). Contrary to popular belief, studies report that a majority of patients prefer, in fact, numerical information to care alone (Hallowell et al. 1997; Wallsten et al. 1993). As indicated above, there are simple ways to make numbers easier to grasp (see also Fagerlin et al. 2007). In this chapter we have addressed two of them: using absolute risks instead of relative risks, and using natural frequencies rather than conditional probabilities. There is more to transparency than can be covered in this chapter. Pictorial, icon, and graphic representations are other formats that can foster insight (Galesic, Garcia-Retamero et al. 2009; Hawley et

**Conclusion**

Despite the growth of information technology, the problem of statistical illiteracy in health is widespread. Patients and doctors alike often lack skills in basic numeracy, and the problem is further aggravated when the benefits and harms of treatment options need to be evaluated or test results assessed. In the literature, statistical illiteracy is often presented as if it were largely a consequence of cognitive limitations (Kahneman et al. 1982). However, if such “probability blindness” were simply caused by faulty mental software, then we would just have to live with a miswired brain. The only remedy taken thus far has come in some form of paternalism; that is, people have been inhibited from making important health decisions and instead nudged toward outcomes determined by others (Thaler and Sunstein 2008).

Most causes of statistical illiteracy that we discussed, however, are found in the external environment. The majority of information directed at patients is either incomplete, nontransparent, or both. Consequently, the easiest and most effective remedy is to use transparent representations, such as absolute risks or natural frequencies, when presenting health information. This should happen in association with a change in school curricula, so that the next generation is equipped to deal with risk and uncertainty. An educated society must know what questions to ask, what information is missing, and how to translate non-transparent statistics into transparent ones. An educated citizenry will drive change in the media and other institutions by exerting pressure to supply complete and transparent information.

We believe that statistical literacy is a necessary precondition for educated citizens in a functioning democracy. Understanding risks and asking critical questions also has the potential to shape the emotional climate in a society and protect its members from external manipulation. In such a society, citizens can develop a better informed and more relaxed attitude toward health.