



Knowing Your

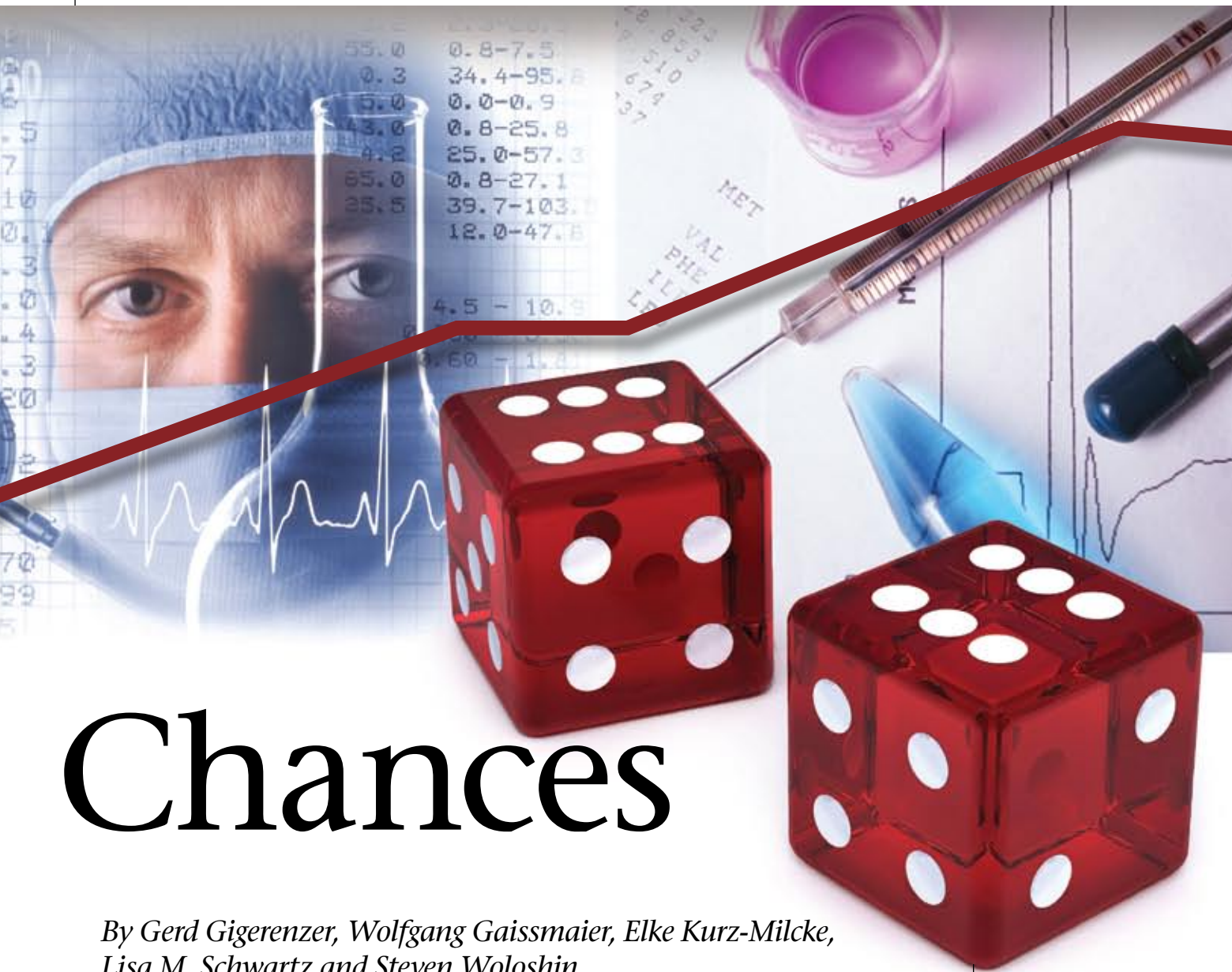
When might a positive HIV test be wrong? Are your chances of surviving cancer better in the U.S. or in England? Learn how to put aside unjustified fears and hopes and how to weigh your real risk of illness—or likelihood of recovery

In a 2007 campaign advertisement, former New York City mayor Rudy Giuliani said, “I had prostate cancer, five, six years ago. My chances of surviving prostate cancer—and thank God, I was cured of it—in the United States? Eighty-two percent. My chances of surviving prostate cancer in England? Only 44 percent under socialized medicine.” Giuliani used these statistics to argue that he was lucky to be living in New York and not in York. This statement was big news. As we will explain, it was also a big mistake.

In 1938 in *World Brain* (Methuen & Co.), English writer H. G. Wells predicted that for an educated citizenship in a modern democracy, statistical thinking would be as indispensable as reading and writing. At the beginning of the 21st century, nearly everyone

living in an industrial society has been taught reading and writing but not statistical thinking—how to understand information about risks and uncertainties in our technological world. That lack of understanding is shared by many physicians, journalists and politicians such as Giuliani who, as a result, spread misconceptions to the public.

Statistical illiteracy is not rooted in inherent intellectual deficits—say, in the lack of a “math gene”—but rather in societal and emotional forces. These influences include the paternalistic nature of the doctor-patient relationship, the illusion of certainty in medicine, and the practice of presenting health information in opaque forms that erroneously suggest big benefits and small harms from interventions. When citizens do not understand the



Chances

By Gerd Gigerenzer, Wolfgang Gaissmaier, Elke Kurz-Milcke, Lisa M. Schwartz and Steven Woloshin

numbers, they are susceptible to political and commercial manipulation of their anxieties and hopes. The result can be serious damage to physical health and emotional well-being.

We show you how to spot three types of statistical manipulation and confusion in medicine, to translate opaque figures into ones that make sense and to use that information to make better medical decisions. To avoid such misunderstandings in the first place, we argue that medical journals, the media and others should communicate risk in more easily understood forms. In addition, we recommend introducing young children to statistical thinking and teaching statistics as a way of solving real-world problems rather than as a purely mathematical discipline.

Trust Your Doctor?

Medicine has held a long-standing antagonism toward statistics. For centuries, treatment was based on an ethic of personal trust as opposed to quantitative facts, which were dismissed as impersonal or irrelevant to the individual. Even today many doctors think of themselves as artists, relying more on intuition and faith in their own judgment than on numbers. For their part, many patients prefer to trust their doctors rather than even asking for data to analyze. For example, in a 2008 unpublished survey by one of us (Gigerenzer) and his colleagues, two thirds of more than 100 American economists said they had not weighed any pros and cons of getting a prostate cancer screening test but simply followed their doctor's recommendation.

Individuals often shy away from statistics because they have an emotional need for certainty.

Moreover, individuals often shy away from statistics because they have an emotional need for certainty—a concept at odds with statistical literacy, which prepares us to make decisions in the face of uncertainty [see box on opposite page]. Much of the public harbors illusory certainty about the reliability of tests such as those for cancers and HIV, suggests a survey Gigerenzer conducted in 2006.

Furthermore, statistically unsophisticated patients and their doctors tend to wildly overestimate the benefits of screening tests and are blind to their harms. For example, mammography reduces the risk of a woman in her 50s dying from breast cancer from about five to four in 1,000 over some 13 years, but 60 percent of a random sample of U.S. women believed the benefit to be 80 times as high. Americans are similarly overenthusiastic about total-body computed tomographic scans: in a random sample of 500 Americans, nearly three quarters said they would prefer a free total-body CT scan to \$1,000 in cash. Yet no professional medical organization endorses such scans, and several discourage them because screening tests such as this one can result in

important harm from a cascade of medical quandries and invasive treatments triggered by ambiguous findings.

A citizen in a modern technological society faces a bewildering array of medical decisions. Should a pregnant woman undergo prenatal screening for chromosomal anomalies at age 35? Should parents send their teenage daughters to be vaccinated against human papillomavirus, to protect them against cervical cancer, despite a few reports that the vaccine (Gardasil) could lead to paralysis? If people want to make informed decisions, they need to understand health statistics. In particular, they need to understand the difference between absolute and relative risks and how to use natural frequencies to infer the true chances of disease from a positive test result. Individuals also should know to trust mortality rates over five-year survival statistics when evaluating screening tests, which look for disease in healthy people. We deal with each of these issues in turn.

Absolute Risks

In October 1995 the U.K. Committee on Safety of Medicines warned that third-generation oral contraceptive pills increased the likelihood of potentially life-threatening blood clots in the legs or lungs twofold—that is, by 100 percent. This information was passed on in “Dear Doctor” letters to 190,000 general practitioners, pharmacists and directors of public health and in an emergency announcement to the media. The news caused great anxiety, and women stopped taking the pill, which led to an estimated 13,000 additional abortions in the following year in England and Wales. For every additional abortion, there was also one extra birth, including some 800 more conceptions among girls younger than 16. (Ironically, abortions and pregnancies are associated with an increased risk of thrombosis that exceeds that of the third-generation pill.)

Such panic could have been avoided had the data been reported in a more straightforward manner. The evidence showed that about one in every 7,000 women who took the second-generation pill had a blood clot; this number increased to two in 7,000 among women who took third-generation pills. That is, the *absolute* risk increase was only one in 7,000 even though the *relative* risk increase was indeed 100 percent. Absolute risks are typically small numbers, whereas the corresponding

FAST FACTS

Know Your Numbers

1» Statistical illiteracy is rooted not in intellectual deficits but in the doctor-patient relationship, the illusion of certainty in medicine, and the practice of presenting health information in opaque forms that erroneously suggest big benefits and small harms from interventions.

2» Without understanding the numbers, citizens are susceptible to political and commercial manipulation of their anxieties and hopes. The result can be serious damage to physical health and emotional well-being.

3» People need to understand the difference between absolute and relative risks and how to use natural frequencies to infer the true chances of disease from a positive test result. Individuals also should know to trust mortality rates over five-year survival statistics when evaluating screening tests.

4» To boost statistical literacy, we also recommend introducing young children to statistical thinking and teaching statistics in school as a way of solving real-world predicaments rather than as a purely mathematical discipline.

Living with Uncertainty

Although people often apply a need for certainty to test results and treatments, no unequivocal answers or absolute cures exist. Risk is unavoidable; it comes with all action or inaction. Here are questions to ask about all risks:

>>1 Risk of what? Understand the outcome to which the risk refers. Is it the risk of dying from a disease, getting the disease or manifesting a symptom?

>>2 What is the time frame? Time frames such as "the next 10 years" are easier to imagine than the widely used "lifetime" risks. They are more informative because risks change over time, and yet such time frames are long enough to enable action.

>>3 How big is the risk? Because there are no zero risks, size is what matters. That number should be expressed in absolute terms—for instance, 13 out of 1,000 50-year-old female smokers die of heart disease within 10 years—or in comparative terms, relating the risk to other ones. For example, a 50-year-old female smoker has about the same chance of dying of heart disease as of lung cancer within the next decade—and these chances are about seven times higher than her risk of perishing in a car accident.

>>4 Does the risk apply to me? Find out if the risk is based on studies of people like you—individuals of your age or sex or with health problems similar to yours.



>>5 What are the harms of "finding out?" Screening tests may lead to false alarms, prompting unnecessary anxiety. When women participate in a 10-year program of annual mammography, every other woman without cancer can expect one or more false-positive test results. Worse, screening tests often detect abnormalities that would never cause symptoms, leading to unnecessary surgery and other invasive treatments.

relative changes tend to look big—particularly when the base rate is low.

Reporting relative risks can create unrealistic hopes as well as undue anxiety. Many patients and doctors evaluate a treatment or test more favorably if benefits are expressed in terms of relative risk reduction. In a 2007 review of experimental studies, for example, psychologist Judith Covey of the University of Durham in England found that when the benefit of a drug was presented in the form of relative risk reduction, 91 percent of Danish general practitioners would recommend it to their patients. But when given the absolute risk reduction, only 63 percent would recommend the same drug.

Information brochures, doctors, medical journals and the media continue to inform the public in terms of relative changes, in part because big numbers make better headlines and generate more attention. One leaflet even conflated the two, stating that hormone replacement therapy (HRT) "has been proven to protect women against colorectal cancer (by up to more than 50 percent)," whereas the risk of breast cancer "may possibly increase by 0.6 percent (six in 1,000)." The data reveal that the 50 percent benefit corresponds to an absolute num-

ber that is less than six in 1,000—meaning that HRT produces more cases of cancer than it prevents. But according to a 2003 study, 60 of 80 women concluded the exact opposite from the leaflet.

Absolute risks are more informative because they take into account information about background rates. Given the absolute risks, a person can derive the relative risks—but not vice versa. After all, a relative risk reduction of 50 percent could describe either a substantial mortality reduction from 200 to 100 in 10,000 patients or a much smaller one from two to one in 10,000 patients. Randomized trials provide some of the best information in med-

This advertisement expresses Lipitor's potency as an impressive-sounding 48 percent relative risk reduction. But the benefit in absolute terms is a far less striking 1.3 percent: after four years, 2.8 percent of patients taking a sugar pill had a stroke as compared with 1.5 percent taking Lipitor.

What most doctors don't know: a positive mammogram may mean a mere one in 10 chance of cancer.

icine, but unless the results are reported adequately, people will not be able to assess them.

Natural Frequencies

Consider a woman who has just received a positive result from a mammogram and asks her doctor: Do I have breast cancer for sure, or what are the chances that I have the disease? In a 2007 continuing education course for gynecologists, Gigerenzer asked 160 of these practitioners to answer that question given the following information about women in the region:

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

What is the best answer to the patient's query?

- A. The probability that she has breast cancer is about 81 percent.
- B. Out of 10 women with a positive mammogram, about nine have breast cancer.
- C. Out of 10 women with a positive mammogram, about one has breast cancer.
- D. The probability that she has breast cancer is about 1 percent.

Gynecologists could derive the answer from the statistics above, or they could simply recall what they should have known anyhow. In either case, the best answer is C; only about one out of every 10 women who test positive in screening actually has breast cancer. The other nine are falsely alarmed. Prior to training, most (60 percent) of the gynecologists answered 90 percent or 81 percent, thus grossly overestimating the probability of cancer. Only 21 percent of physicians picked the best answer—one out of 10.

Many physicians do not know the probabilities that a person has any disease given a positive screening test—that is, the positive predictive value of that test. Nor can they estimate it from conditional probabilities such as the test's sensitivity (the probability of testing positive in the presence of the disease) and the false-positive rate. Such innumeracy causes undue fear. Months after receiving a false-positive mammogram, one in two women reported considerable anxiety about mammograms and breast cancer, and one in four reported that this anxiety affected her daily mood and functioning.

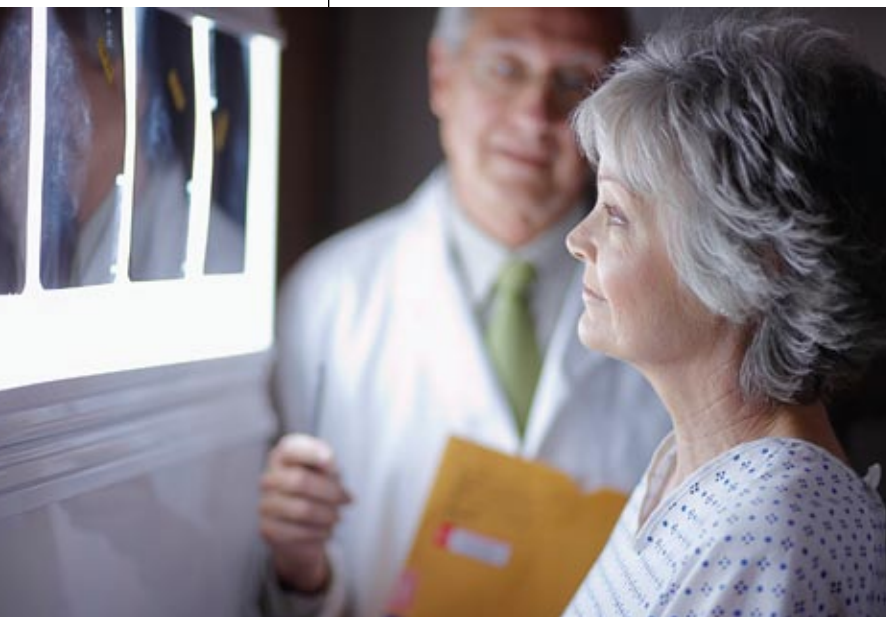
Doctors would more easily be able to derive the correct probabilities if the statistics surrounding the test were presented as natural frequencies. For example:

- Ten out of every 1,000 women have breast cancer.
- Of these 10 women with breast cancer, nine test positive.
- Of the 990 women without cancer, about 89 nonetheless test positive.

Thus, 98 women test positive, but only nine of those actually have the disease. After learning to translate conditional probabilities into natural frequencies, 87 percent of the gynecologists understood that one in 10 is the best answer. Similarly, psychologist Ros Bramwell of the University of Liverpool in England and his colleagues reported in 2006 that only one out of 21 obstetricians could correctly estimate the probability of an unborn child actually having Down syndrome given a positive test. When they were given the relevant natural frequencies, 13 out of 20 obstetricians arrived at the correct answer.

Physicians need to inform patients that no test is perfect, that every test result needs to be interpreted

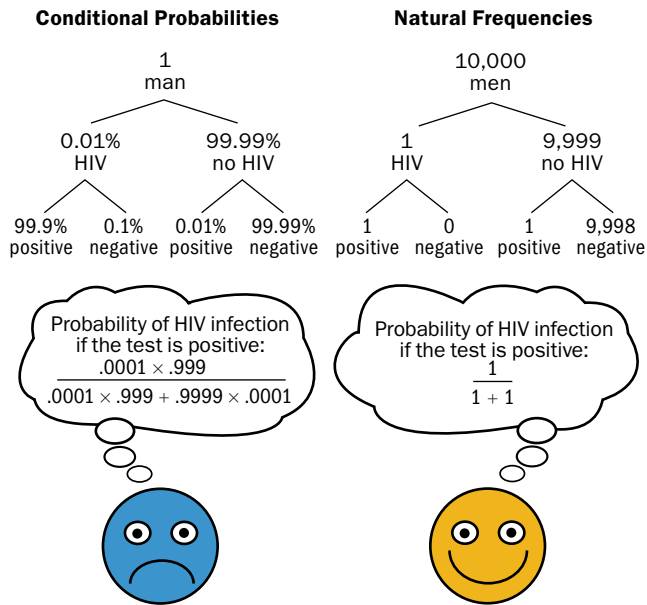
Positive mammograms can cause considerable anxiety. A positive result would be far less alarming, however, if women knew that it might mean a mere 10 percent chance of having breast cancer. In 10 years of testing, one in every two women without cancer can expect at least one positive mammogram.



CORBIS

Forecasting Infection

If your HIV test is positive and you are a man at low risk of infection, what are the chances that you actually harbor the virus? Conditional probabilities (left) leave us with a confusing calculation. Invoking natural frequencies (right) leads to an easy answer: out of every 10,000 men, one is expected to be infected with HIV and will test positive; out of the uninfected men, one should also test positive. Thus, two test positive, and one of these is infected. In other words, your chances of infection given a positive result are not 100 percent; instead they are 50 percent.



with care, or that a test needs to be repeated. Every woman who gets a mammogram should be told that many of the suspicious results are false alarms. A similar uncertainty exists with all such tests, even HIV tests. At a 1987 AIDS conference, then Florida senator Lawton Chiles reported that of 22 blood donors in Florida who had been notified that they had tested positive for HIV, seven committed suicide. Although the test for HIV picks up 99.9 percent of true infections, and 99.99 percent of its negative results are accurate, a very low base rate among low-risk heterosexual men means the chance of infection can be as low as 50 percent when a man tests positive in screening [see box above]. (When the base rate is higher, however, as it is in the case of homosexual men who have unprotected sex or intravenous drug users who share needles, the chance of true infection with a positive HIV test result is almost certain. So the base rate in a population determines the meaning of a positive test result.)

Mortality Matters

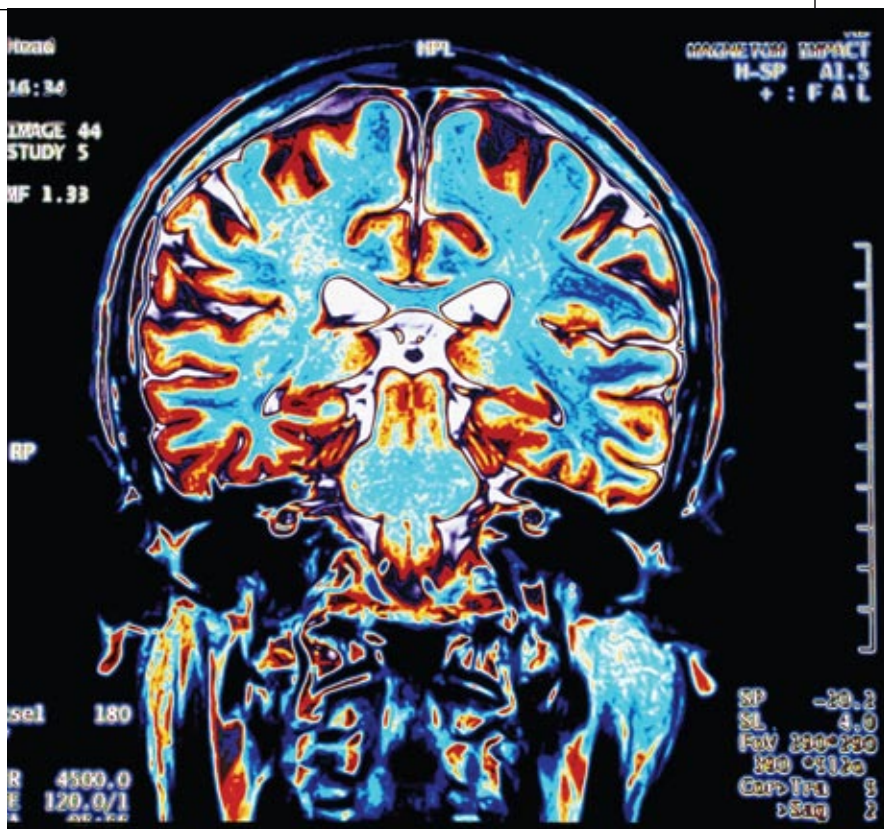
While running for president, Giuliani claimed that health care in the U.S. was superior to that in England. He apparently used data from the year 2000, when 49 British men in every 100,000 were diagnosed with prostate cancer, of whom 28 died within five years—about 44 percent. Using a similar approach, he cited a corresponding 82 percent five-year survival rate in the U.S., suggesting that Amer-

icans with prostate cancer were twice as likely to survive as their British counterparts were. That implication, however, is false because these survival statistics largely reflect diagnostic differences between the two countries rather than better treatment and prolonged survival in the U.S.

To understand why, imagine a group of prostate cancer patients diagnosed (by their symptoms) at age 67 in the U.K., all of whom die at 70. Each survived only three years, so the five-year survival of this group is 0 percent. Now imagine that the same group is diagnosed in the U.S., where doctors detect most prostate cancer by screening for prostate-specific antigens (PSA). (The PSA test is not routinely used in Britain.) These U.S. patients are diagnosed earlier, at age 60, but they all still die at age 70. All have now survived 10 years, and thus their five-year survival rate is 100 percent. Even though the survival rate has changed dramatically,

(The Authors)

GERD GIGERENZER is director at the Max Planck Institute for Human Development and the Harding Center for Risk Literacy in Berlin. He is author of *Gut Feelings: The Intelligence of the Unconscious* (Penguin, 2007). **WOLFGANG GAISSMAIER** studies risk perception as chief research scientist at the Harding Center. **ELKE KURZ-MILCKE** is a psychologist at the Harding Center. **LISA M. SCHWARTZ** and **STEVEN WOLOSHIN** are associate professors of medicine and community and family medicine at the Dartmouth Institute for Health Policy and Clinical Practice.



Magnetic resonance images such as the one above (of a healthy brain) may be used to diagnose brain cancer. Companies and organizations offer such scans as screening tools, suggesting that people need one to be sure they do not have a brain tumor. But brain cancer is quite rare, and there is no evidence that screening is beneficial.

nothing has changed about the time of death. This example shows how setting the time of diagnosis earlier can boost survival rates (lead-time bias), even if no life is prolonged or saved [see illustration on opposite page].

Spuriously high survival rates can also result from overdiagnosis, the detection of abnormalities that are technically cancer but will never progress to cause symptoms in the patient's lifetime. Say 1,000 men with progressive cancer do not undergo screening. After five years 440 are still alive, which results in a survival rate of 44 percent. Meanwhile in another population of men, PSA screening detects 1,000 people with progressive cancer and 2,000 people with nonprogressive cancer (who by definition will not die of cancer in five years). These nonprogressive cases are now added to the 440 who survived progressive cancer, which inflates the survival rate to 81 percent. Although the survival rate changed dramatically, the number of people who die has not changed at all.

In the U.S., screening for prostate cancer using the PSA test in the late 1980s led to an explosion in

the number of new prostate cancer diagnoses. In Britain, the effect has been much smaller because of far less use of the PSA test. This diagnostic disparity largely explains why five-year survival for prostate cancer is higher in the U.S. (The most recent figures are 98 percent five-year survival in the U.S. versus 71 percent in Britain.)

Despite the difference in survival rates, mortality rates in the two countries are close to the same: about 26 prostate cancer deaths per 100,000 American men versus 27 per 100,000 in Britain. That fact suggests the PSA test has needlessly flagged prostate cancer in many American men, resulting in a lot of unnecessary surgery and radiation treatment, which often leads to impotence or incontinence.

Because of overdiagnosis and lead-time bias, changes in five-year survival rates have no reliable relation to changes in mortality when patterns of diagnosis differ. And yet many official agencies continue to talk about five-year survival rates. A recent report by the U.K. Office for National Statistics

noted that five-year survival for colon cancer was 60 percent in the U.S. as compared with 35 percent in Britain. Experts dubbed this finding "disgraceful" and called for a doubling of government spending on cancer treatment. In fact, the mortality rate for colon cancer in Britain is about the same as that in the U.S. In an even stranger case, an ad for the prestigious University of Texas M. D. Anderson Cancer Center conflated survival rates with mortality rates: "as national *mortality* rates for prostate cancer fluctuated between 1960 and 1990, five-year *survival* rates for prostate cancer among M. D. Anderson patients continued to improve" (emphasis added).

Mortality rates are far more reliable indicators of the value of screening programs than are five-year survival rates, which boost survival because of earlier diagnoses and overdiagnoses. So should a man get a PSA test or a smoker undergo a CT scan to screen for lung cancer? Both exams find more early-stage cancers—but neither has been shown to reduce mortality.

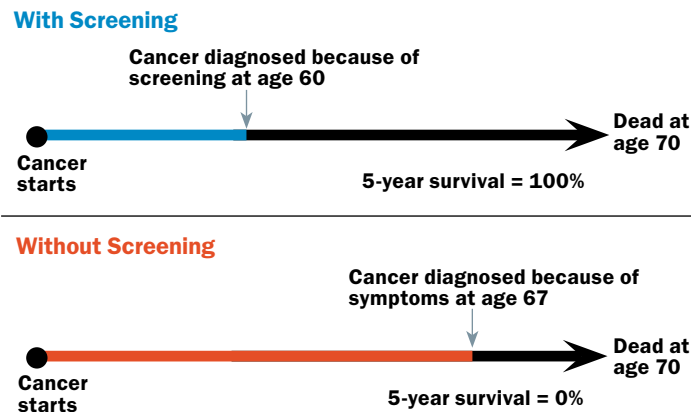
People commonly regard screening as a safeguard for their health, even if an illness is rare. But additional testing may lead to unnecessary medical interventions that can result in harm, which means there is nothing "safe" about this strategy. And for the many overdiagnosed patients, treatment can only cause harm. An epidemic of diagnoses can be as dangerous to our health as disease is.



More Science

See the *Psychological Science in the Public Interest* article, "Helping Doctors and Patients Make Sense of Health Statistics," on which this story is based, at the Association for Psychological Science's Web site: www.psychologicalscience.org

Survival statistics depend on the timing of diagnosis, making such figures misleading. A cancer diagnosis at age 60 (top) can dramatically boost five-year survival rates compared with a diagnosis seven years later (bottom) without changing the time of death.



Solving Problems

Statistical misunderstandings would be far less frequent if researchers, doctors and the media used straightforward figures instead of confusing ones: absolute risks instead of relative risks, natural frequencies instead of conditional probabilities, and mortality rates instead of five-year survival rates. In addition to changing the reporting of health statistics, we need to better educate our young people in the science of risk.

Today the U.S. mathematics curriculum centers on the mathematics of certainty—from arithmetic to calculus—and instruction in probability and statistics occurs too late. As H. G. Wells suggested, statistics should be taught as early as reading and writing. Indeed, the U.S. National Council of Teachers of Mathematics has been pushing educators for years to begin instruction in statistics and probability in primary school. If children learned to deal with an uncertain world in a playful way, much of the collective statistical illiteracy would be history.

Furthermore, teachers need to approach statistics differently. Instead of instructing students about how to apply formulas to toy problems involving cards and dice, teachers should show them how to use numbers to solve real-world predicaments. Statistics might even be wrenched away from math educators to create a problem-solving field connected to teaching health in schools. Such a new field might help young people make better decisions about drugs, alcohol use, driving, biotechnology and other relevant health issues.

In an excellent example of this approach, one secondary school textbook tells the real story of a 26-year-old single mother who tested positive in a routine HIV test, lost her job, moved into a halfway house with other HIV-positive residents, had unprotected sex with one of them, eventually devel-

oped bronchitis, and was asked by her new doctor to take the HIV test again. She did, and the result was negative, as was her original blood sample when it was retested. The woman had lived through a nightmare because her physicians did not realize that a positive test result is not definitive but that instead, in this woman's case, it means just a 50 percent chance of being infected, because she was in a low-risk group.

Statistical literacy can change lives, helping individuals make better personal choices, recognize misleading advertisements and public service messages, and develop a more relaxed attitude toward their health. The dream of statistical literacy embodies the Enlightenment ideal of people's emergence from their self-imposed immaturity. In Immanuel Kant's words, "Dare to know!" **M**

(Further Reading)

- ◆ **Communicating Statistical Information.** Ulrich Hoffrage, Samuel Lindsey, Ralph Hertwig and Gerd Gigerenzer in *Science*, Vol. 290, pages 2261–2262; December 22, 2000.
- ◆ **Calculated Risks: How to Know When Numbers Deceive You.** Gerd Gigerenzer. Simon & Schuster, 2002.
- ◆ **Should I Be Tested for Cancer: Maybe Not and Here's Why.** H. Gilbert Welch. University of California Press, 2004.
- ◆ **Making Numbers Matter: Present and Future Research in Risk Communication.** Angela Fagerlin, Peter A. Ubel, Dylan M. Smith and Brian J. Zikmund-Fisher in *American Journal of Health Behavior*, Vol. 31, Supplement 1, pages S47–S56; September/October 2007.
- ◆ **A Meta-Analysis of the Effects of Presenting Treatment Benefits in Different Formats.** Judith Covey in *Medical Decision Making*, Vol. 27, No. 5, pages 638–654; September/October 2007.
- ◆ **How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics.** Joan Garfield and Dani Ben-Zvi in *International Statistical Review*, Vol. 75, No. 3, pages 372–396; December 2007.
- ◆ **Know Your Chances: Understanding Health Statistics.** Steven Woloshin, Lisa M. Schwartz and H. Gilbert Welch. University of California Press, 2008.
- ◆ **Healthy Skepticism.** Lisa M. Schwartz, Steven Woloshin and H. Gilbert Welch in *Washington Post*. Available at www.vaoutcomes.org/washpost.php