

Helping Physicians Understand Screening Tests Will Improve Health Care

By Gerd Gigerenzer

Medical doctors tend to think of psychologists as therapists, useful for emotionally disturbed patients, but not for members of their own trade. Research on transparent risk communication is beginning to change that view, however.

As a young researcher, I was struck by a study conducted by David Eddy, now Senior Advisor for Health Policy and Management at Kaiser Permanente. He asked American physicians to estimate the probability that a woman had breast cancer given a positive screening mammogram and provided them with the relevant information: a base rate of 1 percent, a sensitivity of 80 percent, and a false-positive rate of 9.6 percent. Approximately 95 out of 100 physicians wrongly reckoned this probability to be around 75 percent, whereas the correct answer is 7.7 percent (Eddy, 1982). Eddy concluded that many physicians make major errors in statistical thinking that threaten the quality of medical care.

How could doctors not have known the answer? Even if some of the doctors tested were “mathematically challenged,” they should already have known that only about one in 10 women with a positive screening mammogram has cancer. Mammography is one of the most frequently tested medical procedures, with widely published results about its accuracy. But most doctors don’t have the time to read medical journals, and few women know that a positive mammogram is like an activated car alarm — usually a false call. As a result, millions of women who test positive every year are unnecessarily frightened.

In a study of German doctors, my colleague Ulrich Hoffrage, University of Lausanne, Switzerland, and I found

Gerd Gigerenzer is Director at the Max Planck Institute for Human Development in Berlin. He has taught at the Universities of Munich, Constance, Salzburg, and Chicago. Recent books include *Simple Heuristics that Make Us Smart* (1999, with Peter Todd et al.), *Adaptive Thinking: Rationality in the Real World* (2000), *Calculated Risks* (2002), and *Gut Feelings: The Intelligence of the Unconscious* (2007). He has been the recipient of many awards, including the American Association for the Advancement of Science Prize for Behavioral Science Research.

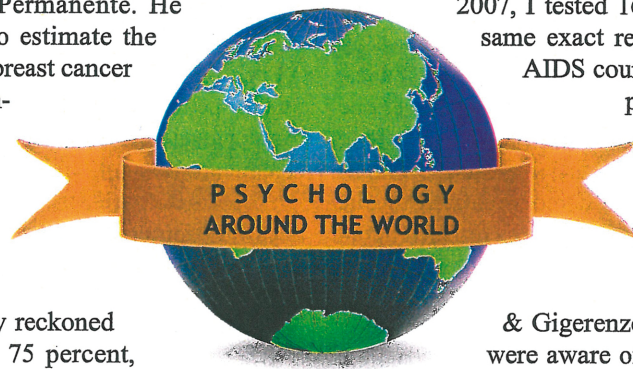
the same problem: Their most frequent estimate was a 90 percent chance of breast cancer given a positive screening mammogram (Hoffrage & Gigerenzer, 1998). In contrast to Eddy’s informal study, our systematic experiments also detected a shocking variability: Physicians’ estimates ranged from 1 percent to 90 percent! Perhaps these physicians simply had no interest in breast cancer statistics? In 2007, I tested 160 gynecologists and found the same exact result. Likewise, the majority of AIDS counselors tested mistakenly took a positive HIV test to be certain (Gigerenzer et al., 1998), and doctors given positive colorectal cancer screening results estimated the probability to be between 1 percent and 99 percent (Hoffrage & Gigerenzer, 1998). Poor patients. If they were aware of this variability, they would be rightly apprehensive. Then again, could *you* solve the Eddy problem?

What can be done? Eddy suggested teaching Bayes’s rule, but in the rare cases where medical schools do so, most students have already forgotten the formula a few weeks after the exam. In 1995, Hoffrage and I found an effective solution: presenting the numbers in natural frequencies instead of conditional probabilities (e.g., sensitivity, false-positive rate). To turn Eddy’s probabilities into natural frequencies, take a sample of 1,000 women. We expect that 10 have cancer, among which eight will test positive; among those without cancer, some 95 will test positive. Now it is easy to see that of the 103 women who test positive, only eight actually have breast cancer — roughly 1 out of 10. The majority of doctors we studied suddenly understood the problem. For those who were slower, a little training helped. For instance, after a one-hour lecture that I delivered to a group of 150 gynecologists, 87 percent understood the chances of cancer given a positive test compared with only 21 percent beforehand.

Another numerical representation that tends to cloud doctors’ minds is relative risk. We read that mammography



This article is one in a series of columns in which leading international psychological scientists share their work and experiences with *Observer* readers.



screening reduces the risk of dying of breast cancer by 25 percent. Many people believe this to mean that the lives of 250 out of 1,000 women are saved, whereas a group of Swiss gynecologists' interpretations varied between one in 1,000 and 750 in 1,000! How large is the actual benefit? Randomized trials showed that, out of 1,000 women not screened, four died of breast cancer within about 10 years, whereas among those who were screened, three died. Thus, the absolute risk reduction is one out of 1,000 women, or 0.1 percent, whereas the relative risk reduction is 25 percent. In a representative 2006 survey of 1,000 German citizens, I found that hardly anyone understands what the 25 percent means. Other sources of confusion are single-event probabilities and five-year survival rates.

Continuing education of doctors has traditionally been in the hands of the pharmaceutical industry, combined with an attractive all-expenses-paid vacation. Doctors are "informed" about the company's products, and for many, this is their main contact with medical developments. Germany and other European countries have stopped this business by law, although it continues to be practiced in the United States. A survey showed that risk communication was number one on doctors' list of interests, and so I've ended up teaching some 1,000 gynecologists over the last two years, a wonderful and moving experience. Some doctors realized for the first time that they were not alone in being confused about percentages and, most importantly, that the problem was not their own mental capabilities but the external representation of information.

Apart from teaching physicians to recognize misleading representations and translate them into transparent ones, we have written about risk communication in local journals that doctors actually read, as well as the top medical journals, which are unfortunately less widely read (Elmore & Gigerenzer, 2005; Gigerenzer & Edwards, 2003; Hoffrage et al., 2000; Mata et al., 2005). We try, with mixed results, to persuade brochures and medical journals to adopt transparent risk communication as a norm, and medical schools to incorporate it into their curricula. On the positive side, the term *natural frequencies* has since entered the vocabulary of evidence-based medicine, and a private sponsor is currently setting up a research center to expand our efforts. At the same time, we often meet fierce resistance from institutions with conflicts of interest or reluctance to rethink their agenda.

Cancer information brochures, for example, provide a great challenge. Most brochures in the United States and in Europe do not present medical research candidly. For instance, benefits are often presented as relative risk reductions, as they are bigger figures, whereas potential harms, if mentioned at all, are reported in less impressive absolute risks, such as that one out of 1,000 women will contract mammogram-induced breast cancer. A few health organizations have responded to our work and reworded their publications accordingly, whereas representatives of others have openly told me that their primary goal is to foster compliance, not comprehension.

Strong resistance to transparency also comes from governments. In response to my book *Calculated Risks* (2002), Karl Lauterbach, then an advisor to the German Minister of Health and now a member of the German parliament, publicly defended the ministry's use of relative risks. He stated that the ministry's responsibility is to inform the general public, not individual women, and that only doctors should inform patients in terms of absolute risks. If this distinction seems as arbitrary to you as it is to me, it might help to know that, with the help of the 25 percent (and even a 35 percent) figure, the German parliament was persuaded to pass a law ensuring women aged 50-69 free access to mammography screening. Yet conflicts of interest are not limited to Germany alone. A few years ago, I presented the program of transparent risk communication to the National Cancer Institute in Bethesda, MD. Two officials took me aside afterwards and lauded the program for its potential to make healthcare more rational. I asked if they intended to implement it. Their answer was "no." And why not? As they explained, transparency in this form was bad news for the government — a benefit of only 0.1 percent instead of 25 percent would make poor headlines for the upcoming election! In addition, their board was staffed by the presidential administration, for whom transparency in health care was not (and is not) a priority.

Informed patients and shared decision-making are beautiful democratic ideals. Yet they will remain science fiction until medical evidence is properly understood and presented. Here, psychology can help both physicians and patients alike. ♦

References

- Eddy, D.M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249-267). Cambridge: Cambridge University Press.
- Elmore, J., & Gigerenzer, G. (2005). Benign breast disease — The risks of communicating risk. *New England Journal of Medicine*, 353, 297-299.
- Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon & Schuster.
- Gigerenzer, G., & Edwards, A.G.K. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal*, 327, 741-744.
- Gigerenzer, G., Hoffrage, U., & Ebert, A. (1998). AIDS counseling for low-risk clients. *AIDS Care*, 10, 197-211.
- Hoffrage, U., & Gigerenzer, G. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704.
- Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, 73, 538-540.
- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290, 2261-2262.
- Mata, J., Dieckmann, A., & Gigerenzer, G. (2005). Transparent risk communication made simple, or How to avoid using confusing probabilities. *Zeitschrift für Allgemeine Medizin*, 81, 537-541.