

Simple Tools for Understanding Risks: From Innumeracy to Insight

Gerd Gigerenzer and Adrian Edwards

Bad presentation of medical statistics such as the risks associated with a particular intervention can lead to patients making poor decisions on treatment. Particularly confusing are single event probabilities, conditional probabilities (such as sensitivity and specificity), and relative risks. How can doctors improve the presentation of statistical information so that patients can make well-informed decisions?

The science fiction writer H. G. Wells predicted that in modern technological societies statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write. How far have we got, a hundred or so years later? A glance at the literature shows a shocking lack of statistical understanding of the outcomes of modern technologies, from standard screening tests for HIV infection to DNA evidence. For instance, doctors, with an average of 14 years of professional experience were asked to imagine using the Haemocult test to screen for colorectal cancer.^{1,2} The prevalence of cancer was 0.3%, the sensitivity of the test was 50%, and the false positive rate was 3%. The doctors were asked: What is the probability that someone who tests positive actually has colorectal cancer? The correct answer is about 5%. However, the doctors' answers ranged from 1% to 99%, with about half of them estimating the probability as 50% (the sensitivity) or 47% (sensitivity minus false positive rate). If patients knew about this degree of variability and statistical innumeracy they would be justly alarmed.

Statistical innumeracy is often attributed to problems inside our minds. We disagree: The problem is not simply internal but lies in the external representation of information, and hence a solution exists. Every piece of statistical information needs a representation—that is, a form. Some forms tend to cloud minds, while others foster insight. We know of no medical institution that leaches the power of statistical representations; even worse, writers of information brochures for the public seem to prefer confusing representations.^{2,3}

Here we deal with three numerical representations that foster confusion: single event probabilities, conditional probabilities, and relative risks. In each case we show alternative representations that promote insight (Table 1). These “mind tools” are simple to learn. Finally, we address questions of the framing (expression) and manipulation of information and how to minimise these effects.

Single-Event Probabilities

The statement “There is a 30% chance of rain tomorrow” is a probability statement about a single event: It will either rain or not rain tomorrow. Single event, probabilities are a steady source of miscommunication because, by definition, they leave open the class of events to which the probability refers. Some people will interpret this statement as meaning that it will rain tomorrow in

*Table 1:
Examples of Confusing Statistical Information, with Alternatives that Foster Insight*

Type of information	Examples	How to foster insight
Single event probabilities	"You have a 30% chance of a side effect from this drug"	Use frequency statements: "Three out of every 10 patients have a side effect from this drug"
Conditional probabilities	The probability of a positive test result if the patient has the disease (sensitivity) The probability of a negative test result if the patient does not have the disease (specificity) The probability of the disease if the patient has a positive test result (positive predictive value)	Use natural frequencies, alone or together with conditional probabilities (see box)
Relative risks	If four out of every 1000 women (age 40 or older) who do not undergo mammography screening die of breast cancer, compared with three out of every 1000 who are screened, the benefit is often presented as a relative risk: "Mammography reduces breast cancer mortality by 25%."	Use absolute risks, alone or together with relative risks: "In every 1000 women who undergo screening one will be saved from dying of breast cancer." Use the number needed to treat or harm: "To prevent one death from breast cancer, 1000 women need to undergo screening for 10 years."

30% of the area, others that it will rain 30% of the time, and a third group that it will rain on 30% of the days like tomorrow. Area, time, and days are examples of reference classes, and each class gives the probability of rain a different meaning.

The same ambiguity occurs in communicating clinical risk, such as the side effects of a drug. A psychiatrist prescribes fluoxetine (Prozac) to his mildly depressed patients. He used to tell them that they have "a 30% to 50% chance, of developing a sexual problem" such as impotence or loss of sexual interest.² Hearing this, patients were anxious. After learning about the ambiguity of single event probabilities, the psychiatrist changed how he communicated risk. He now tells patients that of every ten people who take fluoxetine three to five will experience a sexual problem. Patients who were informed in terms of frequencies were less anxious about taking Prozac. Only then did the psychiatrist realise that he had never checked what his patients had understood by "a 30% to 50% chance of developing a sexual problem." It turned out that many had assumed that in 30% to 50% of their sexual encounters something would go awry. The psychiatrist and his patients had different reference classes in mind: The psychiatrist was thinking in terms of patients, but the patients were thinking in terms of their own sexual encounters.

Frequency statements always specify a reference class (although the statement may not specify it precisely enough). Thus, misunderstanding can be reduced by two mind tools: specifying a reference class before giving a single event probability or only using frequency statements.

Conditional Probabilities

The chance of a test detecting a disease is typically communicated in the form of a conditional probability, the sensitivity of the test: “If a woman has breast cancer the probability that she will have a positive result on mammography is 90%.” This statement is often confused with: “If a woman has a positive result on mammography the probability that she has breast cancer is 90%.” That is, the conditional probability of A given B is confused with that of B given A.⁴ Many doctors have trouble distinguishing between the sensitivity, the specificity, and the positive predictive value of test—three conditional probabilities. Again, the solution lies in the representation.

Consider the question “What is the probability that a woman with a positive mammography result actually has breast cancer?” The box shows two ways to represent the relevant statistical information: in terms of conditional probabilities and natural frequencies. The information is the same (apart from rounding), but with natural frequencies the answer is much easier to work out. Only seven of the 77 women who test positive actually have breast cancer, which is one in 11 (9%). Natural frequencies correspond to the way humans have encountered statistical information during most of their history: They are called “natural” because, unlike conditional probabilities or relative frequencies, they all refer to the same class of observations.⁵ For instance, the natural frequencies “seven women” (with a positive mammogram and cancer) and “70 women” (with a positive mammogram and no breast cancer) both refer to the same class of 1000 women. In contrast, the conditional probability 90% (the sensitivity) refers to the class of eight women with breast cancer, but the conditional probability 7% (the specificity) refers to a different class of 992 women without breast cancer. This switch of reference class can confuse the minds of doctors and patients alike.

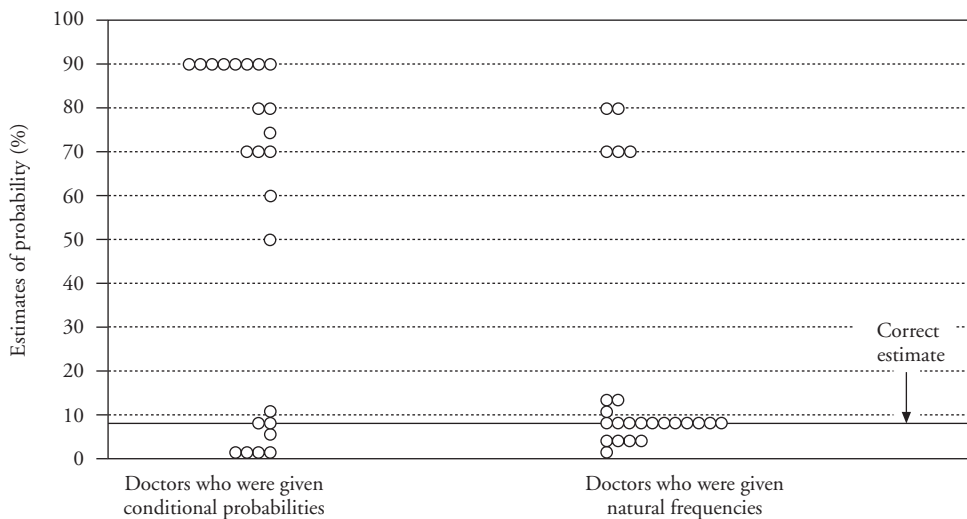


Figure 1. Doctors’ estimates of the probability of breast cancer in women with a positive result on mammography, according to whether the doctors were given the statistical information as conditional probabilities or natural frequencies (each point represents one doctor).²

Figure 1 shows the responses of 48 doctors, whose average professional experience was 14 years, to the information given in the box, except that the statistics were a base rate of cancer of 1%, a sensitivity of 80%, and a false positive rate of 10%.^{1 2} Half the doctors received the information in conditional probabilities and half in natural frequencies. When asked to estimate the probability that a woman with a positive result actually had breast cancer, doctors who received conditional probabilities gave answers that ranged from 1% to 90%, and very few gave the correct answer of about 8%. In contrast most doctors who were given natural frequencies gave the correct answer or were close to it. Simply stating the information in natural frequencies turned much of the doctors' innumeracy into insight, helping them understand the implications of a positive result as it would arise in practice. Presenting information in natural frequencies is a simple and effective mind tool to reduce the confusion resulting from conditional probabilities.⁶ This is not the end of the story regarding the communication of risk (which requires adequate exploration of the implications of the risk for the patient concerned, as described elsewhere in this issue⁷), but it is an essential foundation.

Relative Risks

Women aged over 50 years are told that undergoing mammography screening reduces their risk of dying from breast cancer by 25%. Women in high risk groups are told that bilateral prophylactic mastectomy reduces their risk of dying from breast cancer by 80%.⁸ These numbers are relative risk reductions. The confusion produced by relative risks has received more attention in the medical literature than that of single event or conditional probabilities.^{9 10} Nevertheless, few patients realise that the impressive 25% figure means an absolute risk reduction of only one in 1000: of 1000 women who do not undergo mammography about four will die from breast cancer within 10 years, whereas out of 1000 women who do three will die.¹¹ Similarly, the 80% figure for prophylactic mastectomy refers to an absolute risk reduction of four in 100: five in 100 women in the high risk group who do not undergo prophylactic mastectomy will die of breast cancer, compared with one in 100 women who have had a mastectomy. One reason why most women misunderstand relative risks is that they think that the number relates to women like themselves

Two ways of representing the same statistical information

Conditional probabilities

The probability that a woman has breast cancer is 0.8%. If she has breast cancer, the probability that a mammogram will show a positive result is 90%. If a woman does not have breast cancer the probability of a positive result is 7%. Take, for example, a woman who has a positive result. What is the probability that she actually has breast cancer?

Natural frequencies

Eight out of every 1000 women have breast cancer. Of these eight women with breast cancer seven will have a positive result on mammography. Of the 992 women who do not have breast cancer some 70 will still have a positive mammogram. Take, for example, a sample of women who have positive mammograms. How many of these women actually have breast cancer?

who take part in screening or who are in a high risk group. But relative risks relate to a different class of women: to women who die of breast cancer without having been screened.

Confusion caused by relative risks can be avoided by using absolute risks (such as one in 1000) or the number needed to treat or to be screened to save one life (the NNT, which is the reciprocal of the absolute risk reduction and is thus essentially the same representation as the absolute risk). However, health agencies typically inform the public in the form of relative risks.^{2 3} Health authorities tend not to encourage transparent representations and have themselves sometimes shown innumeracy, for example when funding proposals that report benefits in relative rather than absolute risks because the numbers look larger.¹² For authorities that make decisions on allocation of resources the population impact number (the number of people in the population among whom one event will be prevented by an intervention) is a better means of putting risk into perspective.¹³

The Reference Class

In all these representations the ultimate source of confusion or insight is the reference class. Single event probabilities leave the reference class open to interpretation. Conditional probabilities such as sensitivity and specificity refer to different classes (the class of people with and without illness, respectively), which makes their mental combination difficult. Relative risks often refer to reference classes that differ from those to which the patient belongs, such as the class of patients who die of cancer rather than those who participate in screening. Using transparent representations such as natural frequencies clarifies the reference class.

Framing

Framing is the expression of logically equivalent information (whether numerical or verbal) in different ways.¹⁴ Studies of the effects of verbal framing on interpretation and decision making initially focused on positive versus negative framing and on gain versus loss framing.¹⁵ Positive and negative frames refer to whether an outcome is described, for example, as a 97% chance of survival (positive) or a 3% chance of dying (negative). The evidence is that positive framing is more effective than negative framing in persuading people to take risky treatment options.^{16 17} However, gain or loss framing is perhaps even more relevant to communicating clinical risk, as it concerns the implications of accepting or declining tests. Loss framing considers the potential losses from not having a test, such as, in the case of mammography, loss of good health, longevity, and family relationships. Loss framing seems to influence the uptake of screening more than gain framing (the gains from taking a test, such as maintenance of good health).¹⁸

Visual representations may substantially improve comprehension of risk.¹⁹ They may enhance the time efficiency of consultations. Doctors should use a range of pictorial representations (graphs, population figures) to match the type of risk information that the patient most easily understands.²⁰

Manipulation

It may not seem to matter whether the glass is half full or half empty, yet different methods of presenting risk information can have important effects on outcomes among patients. That verbal and statistical information can be presented in two or more ways means that an institution or screening programme may choose the one that best serves its interests. For instance, a group of gynaecologists informed patients in a leaflet, of the benefits of hormone replacement therapy in terms of relative risk (large numbers) and of harms in absolute risk (small numbers).²

Pictorial representations of risk are not immune to manipulation either. For example, different formats such as bar charts and population crowd figures could be used.²¹ Or the representation could appear to support short term benefits from one treatment rather than long term benefits from another.²² Furthermore, within the same format, changing the reference class may produce greatly differing perspectives on a risk and may thus affect patients' decisions. Figure 2 relates to the effect of treatment with aspirin and warfarin in patients with atrial fibrillation. On the left side of the figure the effect of treatment on a particular event (stroke or bleeding) is shown relative to the class of people who have not had the treatment (as in relative risk reduction). On the right side the patient can see the treatment effect relative to a class of 100 untreated people who have not had a stroke or bleeding (as in absolute risk reduction).

The wide scope for manipulating representations of statistical information is a challenge to the ideal of informed consent.^{2 16} Where there is a risk of influencing outcomes and decisions

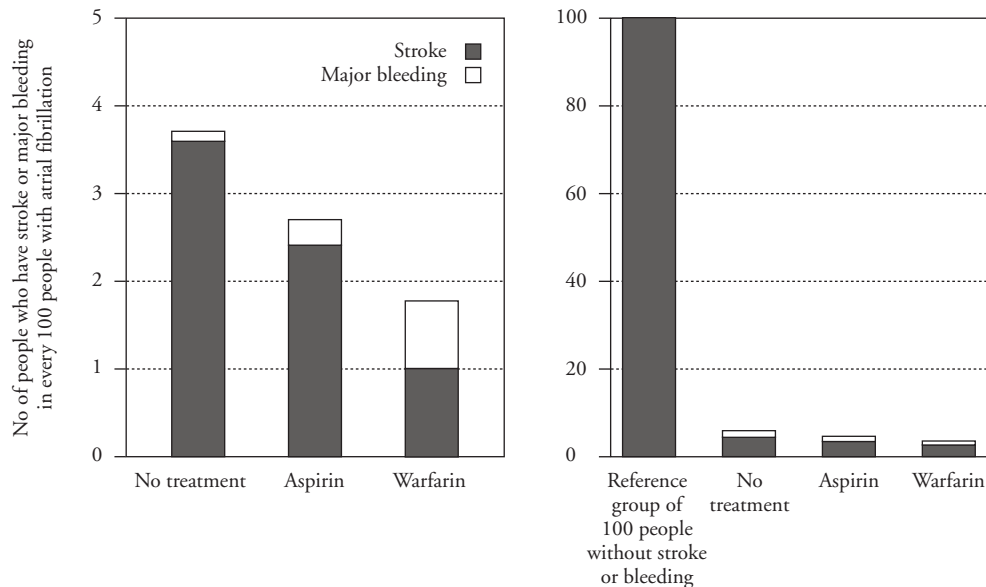


Figure 2. Different representations of the same benefits of treatment: the reduction after treatment in the number of people who have a stroke or major bleeding looks much larger on the left, where the reference class of 100 patients who have not had a stroke or bleeding is not shown.

Summary points

The inability to understand statistical information is not a mental deficiency of doctors or patients but is largely due to the poor presentation of the information.

Poorly presented statistical information may cause erroneous communication of risks, with serious consequences.

Single event probabilities, conditional probabilities (such as sensitivity and specificity) and relative risks are confusing because they make it difficult to understand what class of events a probability or percentage refers to.

For each confusing representation there is at least one alternative, such as natural frequency statements, which always specify a reference class and therefore avoid confusion, fostering insight.

Simple representations of risk can help professionals and patients move from innumeracy to insight and make consultations more time efficient.

Instruction in efficient communication of statistical information should be part of medical curriculums and doctors' continuing education.

among patients, professionals should consistently use representations that foster insight and should balance the use of verbal expressions—for example, both positive and negative frames or both gain and loss frames.

Conclusions

The dangers of patients being misled or making uninformed decisions in health care are countless. One of the reasons is the prevalence of poor representations. Such confusion can be reduced or eliminated with simple mind tools.^{2 23} Human beings have evolved into good intuitive statisticians and can gain insight, but only when information is presented simply and effectively.²⁴ This insight is then the platform for informed discussion about the significance and burden of risks and the implications for the individual or family concerned. It also makes the explanation of diseases and their treatment easier. Instruction in the efficient communication of statistical information should be part of medical curriculums and continuing education for doctors.

Contributors and sources: The research on statistical representations was initially funded by the Max Planck Society and has been published in scientific journals as well summarised in Gerd Gigerenzer's book *Reckoning With Risk: Learning to Live With Uncertainty*. The work on framing is based on research by Adrian Edwards.

Competing interests: None declared.

- 1 Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Acad Med*, 73, 538–540.
- 2 Gigerenzer, G. (2002). *Reckoning with risk: Learning to live with uncertainty*. London: Penguin. (Published in the United States as *Calculated risks: how to know when numbers deceive you*. New York: Simon & Schuster, 2002.)
- 3 Slaytor, E. K., & Ward, J. E. (1998). How risks of breast cancer and benefits of screening are communicated to women: Analysis of 58 pamphlets. *BMJ*, 317, 263–264.
- 4 Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratory results. *N Engl J Med*, 299, 999–1000.
- 5 Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psych Rev*, 102, 684–704.
- 6 Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290 (22), 61–62.
- 7 Paling, J. (2003). Strategies to help patients understand risks. *BMJ*, 327, 745–748.
- 8 Hartmann, L. C., Schaid, D. J., Woods, J. E., Crotty, T. P., Myers, J. L., Arnold, P. G., et al. (1999). Efficacy of bilateral prophylactic mastectomy in women with a family history of breast cancer. *N Engl J Med*, 340, 77–84.
- 9 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.
- 10 Wright, P. (2001). Document-based decision-making. In J. F. Rouet, A. Levonen, & A. Biardeau (Eds.), *Multimedia learning: Cognitive and instructional issues* (pp. 31–43). Amsterdam: Pergamon.
- 11 Nyström, L., Larsson, L.-G., Wall, S., Rutqvist, L., Andersson, I., Bjurstam, N., et al. (1996). An overview of the Swedish randomised mammography trials: Total mortality pattern and the representativity of the study cohorts. *J Med Screen*, 3, 85–87.
- 12 Fahey, T., Griffiths, S., & Peters, T. J. (1995). Evidence based purchasing: Understanding results of clinical trials and systematic reviews. *BMJ*, 311, 1056–1059.
- 13 Heller, R. F., & Dobson, A. J. (2000). Disease impact number and population impact number: Population perspectives to measures of risk and benefit. *BMJ*, 321, 950–953.
- 14 Wilson, D. K., Purdon, S. E., & Wallston, K. A. (1988). Compliance to health recommendations: A theoretical overview of message framing. *Health Educ Res*, 3, 161–171.
- 15 Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- 16 Edwards, A. G. K., Elwyn, G. J., Covey, J., Mathews, 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.
- 17 Kühberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organ Behav Hum Decis Process*, 75, 23–55.
- 18 Edwards, A. G. K., Elwyn, G. J., & Mulley, A. (2002). Explaining risks: Turning numerical data into meaningful pictures. *BMJ*, 324, 827–830.
- 19 Lipkus, I. M., & Hollands, J. G. (1999). The visual communication of risk. *J Natl Cancer Inst Monogr*, 25, 149–162.
- 20 Woloshin, S., & Schwartz, L. (1999). How can we help people make sense of medical data? *Eff Clin Prac*, 2, 176–183.
- 21 Schapira, M., Nattinger, A., & McHorney, C. (2001). Frequency or probability? A qualitative study of risk communication formats used in health care. *Med Decis Making*, 21, 459–467.
- 22 Mazur, D., & Hickam, D. H. (1994). Assumptions patients make when interpreting graphical displays of surgical data. *Theor Surg*, 9, 129–133.
- 23 Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- 24 Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58, 1–73.