BMJ

BMJ 2011;343:d6386 doi: 10.1136/bmj.d6386

OBSERVATIONS

THE ART OF RISK COMMUNICATION

What are natural frequencies?

Doctors need to find better ways to communicate risk to patients

Gerd Gigerenzer *director, Centre for Adaptive Behaviour and Cognition, Max Planck Institute for Human Development, Berlin*

A 2011 Cochrane Review concluded that health professionals and consumers "understood natural frequencies better than probabilities."[1] A 2011 *Annals of Internal Medicine* article reported the opposite, that "natural frequencies are not the best format for communicating the absolute benefits and harms of treatment"[2] How should physicians deal with these contradictory messages?

As is often the case, the contradiction lies in the definitions, not in the data. Ulrich Hoffrage and I introduced the term "natural frequencies" in the late 1990s and conducted the first studies showing that they foster understanding of the positive predictive value among lay people, doctors, and medical students.[3-6] What is a natural frequency? It is a joint frequency of two events, such as the number of patients with disease and who have a positive test result, and is an alternative to presenting the same information in conditional probabilities, such as sensitivities and specificities. Conditional probabilities tend to cloud the minds of many people, including health professionals, as the following problem illustrates (for convenience, probabilities are expressed in percentages).

Assume you use mammography in a certain region to screen for breast cancer. The following information is known:

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

A woman tests positive. What is the chance that she actually has breast cancer?

When I asked 160 gynaecologists this question at the beginning of a continuing medical education session on risk literacy, a majority (60%) believed that the answer was 80% to 90% and 19% believed it to be 1%.[7] If patients knew about this variability, they would rightly be scared.

Natural frequencies, in contrast, help to see through the mental fog. To express them, you take a large enough number of people (such as 100 or 1000, depending on the prevalence) and break this number down into natural frequencies:

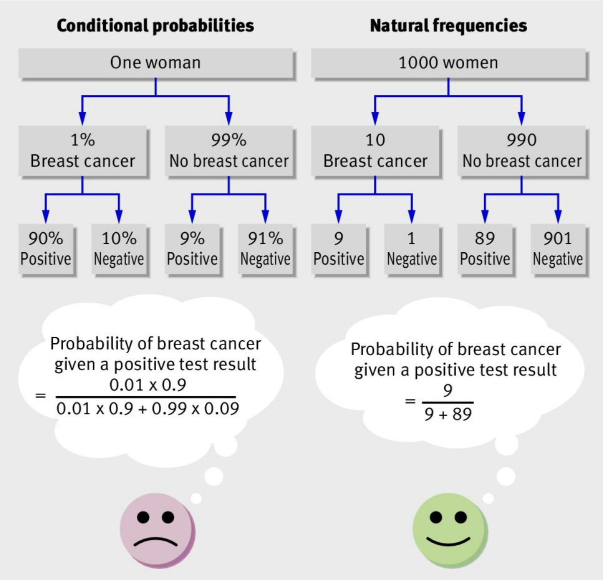
- Ten in every 1000 women have breast cancer.
- Of these 10 women with breast cancer, nine will test positive.
- Of the 990 women without breast cancer, about 89 nevertheless test positive.

Now it is easier to recognise the answer. We expect that 98 women test positive altogether, nine of whom have breast cancer. Therefore, the positive predictive value is 9/98 (9.2%), or roughly one in 10. That is, we expect that of 10 women who test positive nine have false positive results. After just this one session, most of the gynaecologists (87%) had mastered translating sensitivities and false alarm rates into natural frequencies and calculating the positive predictive value.

Natural frequencies facilitate insight. Even 10 year olds can determine the positive predictive value when given natural frequencies, often to the astonishment of their teachers, but are helpless when given conditional probabilities.[8] Why is that? To compute the probability of cancer given a positive test result from probability information (fig, left), we need to use Bayes's rule, a rather complex formula that entails three multiplications. Note that the four conditional probabilities at the bottom of the tree (left) do not add up to 100%. This is because they are normalised with respect to the prevalence of cancer or no cancer. By normalising, the information about prevalence gets lost and then needs to be reintroduced by multiplying it by each conditional probability. Natural frequencies (fig, right), in contrast, do not require these multiplications. The four natural frequencies at the bottom of the tree (right) are not normalised but add up to the total number on the top, simplifying the use of Bayes's rule.

gigerenzer@mpib-berlin.mpg.de

For personal use only: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe



Conditional probabilities (left) make it difficult to infer the positive predictive value, while natural frequencies (right) make it easy. The reason is that the representation simplifies the computation. Both formulas are versions of Bayes's rule. The four probabilities (expressed in percentages) at the bottom of the left tree are conditional probabilities; the four frequencies at the bottom of the right tree are natural frequencies. Natural frequencies are a form of intuitive statistics, corresponding to the way humans encountered information before probability theory was invented

So why did the *Annals of Medicine* article not find this advantage? Simply because it did not test natural frequencies. As shown in the figure, natural frequencies are joint frequencies, such as the number of women (nine) who test positive and who have breast cancer. These differ from simple frequencies, such as two in 10 people who test positive. Similarly, conditional probabilities and simple probabilities are not the same. What the *Annals of Medicine* article did was compare simple percentages (such as 2% of people who took a drug had diarrhoea) and other formats against simple frequencies (20 in every 1000 people who took the drug had diarrhoea), which it called natural frequencies. However, the computational advantage does not apply to simple frequencies.

It does not matter to a computer program whether the input is conditional probabilities or natural frequencies, but to a human being it clearly does. Healthcare providers need to know how to represent information so that their patients can actually understand what it means. Providing a helpful representation is a key skill in the art of communication of risk. Yet understanding test results is not a forte of most doctors themselves. To remedy the situation, every medical curriculum should teach an understanding of health statistics. This is not the same as teaching biostatistics (which corresponds to the left of the figure). For some people, natural frequency trees are ideal; for others, natural frequency grids that represent individuals by icons are best. Effective representations such as these are indeed available, but we need to teach them.

1 Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F. Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database Syst Rev* 2011;(4):CD006776.
2 Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: a randomized trial. *Ann Intern Med* 2011;155:87-96.
3 Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: frequency formats. *Psychol Rev* 1995;102:684-704.
4 Gigerenzer G, Hoffrage U. Overcoming difficulties in Bayesian reasoning: a reply to Lewis and Keren (1999) and Mellers and McGraw (1999). *Psychol Rev* 1999;106:425-30.
5 Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Acad Med* 1998;73:538-40.
6 Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G. Communicating statistical information. *Science* 2000;290:2261-2.
7 Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz, LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest* 2007;8:53-96.
8 Zhu L, Gigerenzer G. Children can solve Bayesian problems: the role of representation in mental computation. *Cognition* 2006;98:287-308.

Cite this as: *BMJ* 2011;343:d6386

© BMJ Publishing Group Ltd 2011