POLICY FORUM: MEDICINE

Communicating Statistical Information

Ulrich Hoffrage,* Samuel Lindsey, Ralph Hertwig, Gerd Gigerenzer

ecisions based on statistical information can mean the difference between life and death—for instance, when a cancer patient has to decide whether to undergo a painful medical procedure based on the likelihood that it will succeed, or when a jury has to decide whether to convict someone based on DNA evidence. Unfortunately, most of us, experts included, have difficulty understanding and combining statistical information effectively.

For example, faculty, staff, and students at Harvard Medical School were asked to estimate the probability of a disease given the following information (1): "If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5 per cent, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs?" The estimates varied wildly, ranging from the most frequent estimate, 95% (given by 27 out of 60 participants), to the correct answer, 2% (given by 11 out of 60 participants) (2). In a study requiring interpretation of mammography outcomes (3), almost all physicians confused the sensitivity of the test (the proportion of positive test results among people with the disease) with its positive predictive value (the proportion of people with the disease among those who receive a positive test result). This is a common confusion that even crops up in scholarly articles (3) and statistical textbooks (4) and certainly affects the ability of lay people (5) to understand the statistical information. Recent discussions of genetic testing have indicated that genetic counselors are experiencing the same difficulty (6).

It makes little mathematical difference whether statistics are expressed as probabilities, percentages, or absolute frequen-

This Policy Forum was collaboratively written to combine work submitted to *Science* independently by the first and second authors.

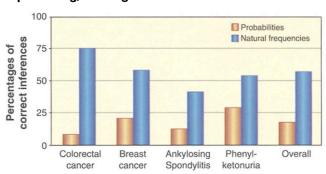


Fig. 1. Interpreting statistics. Medical students' percentage of correct inferences in four realistic diagnostic tasks.

cies. It does, however, make a psychological difference. More specifically, statistics expressed as natural frequencies improve the statistical thinking of experts and non-experts alike.

Natural Frequencies

To illustrate how natural frequencies differ from probabilities, we use the example of a cancer screening test. The probability of col-

orectal cancer can be given as 0.3% [base rate]. If a person has colorectal cancer, the probability that the hemoccult test is positive is 50% [sensitivity]. If a person does not have colorectal cancer. the probability that he still tests positive is 3% [false-positive rate]. What is the probability that a person who tests positive actually has colorectal cancer? A restatement of the same problem in terms of

anatural frequencies would be that out of every 10,000 people, 30 have colorectal cancer. Of these, 15 will have a positive hemoccult test. Out of the remaining 9970 people without colorectal cancer, 300 will still test positive. How many of those who test positive actually have colorectal cancer?

Only 1 out of 24 physicians gave the correct answer when the statistical information was expressed in probabilities (7). When it was presented in natural frequencies, 16 out of 24 other physicians gave the

correct answer: 15 out of 315 (i.e., 5%). Whereas natural frequencies seem to help people make statistical inferences, probabilities apparently hinder them. Unfortunately, in contexts in which the positive predictive value of a test is at issue, statistics are typically expressed and communicated in the form of probabilities, although they can easily be translated into natural

frequencies, as follows:

- 1. Select a population and use the base rate to determine how many people in the population have the disease.
- 2. Take that result and use the test's sensitivity to determine how many people have the disease and a positive test.
- 3. Take the remaining number of healthy people and use the test's false-positive rate to determine how many

people do not have the disease but still test positive.

4. Compare the number obtained in step 2 with the sum of those obtained in steps 2 and 3 to determine how many people with a positive test actually have the disease.

Natural frequencies facilitate inferences because they carry implicit information about base rates and reduce the num-

ber of computations required to determine the positive predictive value of a test (8, 9, 10). They also correspond to the way in which humans have experienced statistical information over most of their history.

Applications in Medicine

To illustrate the effect of natural frequencies, we asked 96 advanced medical students to solve four realistic di-

agnostic tasks. Each participant worked on two probability and two frequency versions; the order of representation format and which task was in which format was balanced (11). For each of the tasks, more participants correctly inferred the likelihood of having the disease given a positive test when the statistics were communicated as natural frequencies (Fig. 1).

Other medical practitioners could also profit from representing statistical information in terms of natural frequencies.

"...statistics
expressed as natural
frequencies improve
the statistical
thinking of experts
and nonexperts
alike."

U. Hoffrage, R. Hertwig, and G. Gigerenzer are at the Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. S. Lindsey is at the Department of Psychology, 102 Gilmer Hall, University of Virginia, Charlottesville, VA 22903, USA.

^{*}To whom correspondence should be addressed. E-mail: hoffrage@mpib-berlin.mpg.de

SCIENCE'S COMPASS

Consider the statistics AIDS counselors must understand and communicate. In Germany, the prevalence of HIV in heterosexual men who are not in any known risk group is around 0.01%. The false-positive rate of the HIV test (in which one blood sample is subjected to multiple tests) is around 0.01%, and its sensitivity is around 99.9% [exact estimates vary (12)]. To ex-

plore how counselors actually communicate these risks, we sent a male, low-risk client to 20 German public. health centers to have 20 HIV tests. During the mandatory pretest counseling, the client asked the counselor about the prevalence, sensitivity, false-positive rate, and the chance that he actually had the virus if the test were positive (13). Not a single counselor communicated the

risks to the client in natural frequencies. Instead, they used probabilities and percentages, and, in the majority of the counseling sessions, the information was either inconsistent or wrong. For instance, one counselor estimated the base rate and the false-positive rate to be around 0.1%, and the sensitivity to be 99.9%, and then stated that the client's probability of infection given a positive test is also 99.9% (applying steps 1 to 4 above to his estimates yields a probability of 50%). In fact, 15 out of the 20 counselors told this low-risk client that it is 99.9% or 100% certain that he has HIV if he tests positive (applying steps 1 to 4 to the numbers found in the literature yields an actual probability of 50%).

Percentages can mislead in other ways. For example, it may sound impressive to learn that mammography screening can reduce the risk of breast cancer fatality in women by 25% [for 50- to 74-year-old women (14)]. However, this percentage does not say anything about the actual frequencies. If 4 out of 1000 women without symptoms die of breast cancer within the next 10 years (15), the relative risk reduction of 25% means that 1 woman in 1000 women who undergo screening would be saved. A woman without symptoms is most likely not one of the 4 to whom the risk reduction applies, but one of the other 996 instead—and many of these women may suffer as a result from the screening. For instance, false-positives occur and, moreover, cancers that grow so slowly that they present little risk will be diagnosed and unnecessarily treated. As long

as health organizations inform women in terms of probabilities and relative risk reduction about the benefits and harms of screening, a truly informed decision is unlikely.

Applications in Law

Determinations of facts and verdicts in legal proceedings often depend on scientific evidence. The communication of statistics

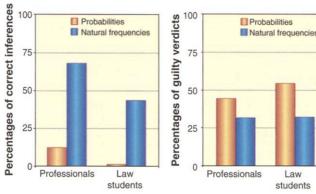


Fig. 2. Interpreting statistics. Legal experts' percentage of correct inferences (left) and of guilty verdicts (right) in two criminal court case files.

is as important to the making of legal decisions by judges, attorneys, forensic experts, and jurors as it is to medical decision-makers (16, 17). In considering the admissibility standards for scientific evidence, the U.S. Supreme Court has specifically indicated that courts need to consider "known or potential rate of error, and the existence and maintenance of standards controlling the technique's operation" (18).

In a study conducted in Germany, we asked 27 professionals who would soon qualify as judges and 127 advanced law students to evaluate two criminal-court case files involving rape (19). In both cases, a DNA match was reported between a DNA sample from the defendant and one recovered from the victim. Aside from this evidence, there was little reason to suspect that the defendant was the perpetrator. Expert testimony reported the frequency of the recovered DNA profile as 1 in 1,000,000 and then stated that it was practically certain that the analysis would show a match for a person who indeed had the DNA profile (in other words, sensitivity = 100%). The expert also reported the rates of technical and human mishaps leading to false-positive results in laboratory tests to be about 0.003 (20).

When these statistics were expressed as probabilities, only 13% of the professionals and under 1% of the law students correctly inferred the probability that the defendant was actually the source of the trace. But when the identical statistics were stated as natural frequencies, 68% and 44% of these same participants made the correct inference (Fig. 2, left). The different ways of expressing the same statistical information altered the verdicts in each case. When the information was presented as probabilities, 45% of the professionals and 55% of the students rendered a verdict of guilty, but only 32% and 33% did so when the same statistics were expressed as natural frequencies (Fig. 2, right). When verdicts hinge on statistical evidence, understanding that evidence is crucial, and pursuing this simple method of fostering statistical insight could contribute to that goal (21, 22).

Implications for Teaching

The beneficial effects of natural frequencies on statistical reasoning in the studies reported above occurred without training or instruction. Systematic training in the use of natural frequencies can even help people to reason with probabilities. The key is to teach representations rather than rules—that is to teach people how to translate probabilities into natural frequencies, as shown in steps 1 to 4. Traditionally, however, students are instead taught how to plug probabilities into mathematical formulas such as Bayes's rule.

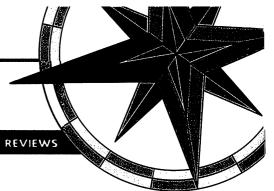
Teaching representations rather than rules-and expressing statistical information in natural frequencies where appropriate—can help to foster the statistical reasoning needed to make sound decisions.

References and Notes

- 1. W. Casscells, A. Schoenberger, T. Grayboys, N. Engl. J. Med. 299, 999 (1978)
- 2. The correct solution can be obtained by applying Bayes's rule.
- 3. D. M. Eddy, in Judgment Under Uncertainty: Heuristics and Biases, D. Kahneman, P. Slovic, A. Tversky, Eds. (Cambridge Univ. Press, Cambridge, 1982), pp. 249-267
- G. Gigerenzer, in A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues, G. Keren, C. Lewis, Eds. (Erlbaum, Hillsdale, NJ, 1993), pp. 313–339.
- 5. J. J. Koehler, Behav. Brain Sci. 19, 1 (1996).
- 6. R. Weiss, Washington Post, 2 December 2000, p. A10. 7. U. Hoffrage, G. Gigerenzer, Acad. Med. 73, 538
- 8. G. Gigerenzer, U. Hoffrage, Psychol. Rev. 102, 684 (1995).
- , Psychol. Rev. **106**, 425 (1999).
- 10. B. Mellers, A. McGraw, *Psychol. Rev.* **106**, 417 (1999). 11. The tasks are displayed at http://www-abc.mpib-
- berlin.mpg.de/users/hoffrage/ papers/4tasks.html. 12. G. J. Stine, Acquired Immune Deficiency Syndrome:
- Biological, Medical, Social, and Legal Issues (Prentice-Hall, Englewood Cliffs, NJ, 1996)
- 13. G. Gigerenzer, U. Hoffrage, A. Ebert, AIDS Care 10, 197 (1998).
- 14. K. Kerlikowske et al., JAMA 273, 149 (1995).
- L. Nystöm, et al., J. Med. Screen. 3, 85 (1997).
 See, e.g., J.J. Koehler, Jurimetrics 34, 21 (1993)
- _, Univ. Colorado Law Rev. **67,** 859 (1996).
- 18. Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993), pp. 593-594.
- 19. S. Lindsey, R. Hertwig, G. Gigerenzer, in preparation.
- 20. An average of estimates based on laboratory proficiency tests. See J. J. Koehler, A. Chia, S. Lindsey, Jurimetrics 35, 201 (1995).
- 21. S. Breyer, Science 280, 537 (1998).
- 22. R. Hertwig, U. Hoffrage, in Frequency Processing and Cognition, P. Sedlmeier, T. Betsch, Eds. (Oxford Univ. Press, New York, in press).
- 23. We thank the German Research Foundation for financial support (Ho 1847/1 and He 2768/6-1).

COMPASS

LETTERS SCIENCE & SOCIETY POLICY FORUM BOOKS ET AL. PERSPECTIVES REVIEWS



Continuing Research into Gulf War Illness

EVIDENCE SUPPORTS A LINK BETWEEN lasting health problems in some Vietnam veterans and wartime exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (2,3,7,8-TCDD) in the defoliant Agent Orange. Martin Enserink, in his News Focus article "Gulf War Illness: The Battle Continues" (2 Feb., p. 812), omits this point in his review of postwar syndromes of the past.



A U.S. Air Force C-123 flies over South Vietnam in 1966, spraying defoliants containing dioxins.

The Agent Orange example suggests that sustained research might yield important information on the cause of health effects observed in Gulf War veterans. A major research effort was initiated in response to health effects observed in people who had contact with Agent Orange in Vietnam. The Environmental Protection Agency and others built on this work and found that 2,3,7,8-TCDD and related compounds ("dioxins") may affect the health of people in the general population (1). It is now feasible to avoid future health problems by preventing this pollution (2), and a global treaty seeking this goal is in the works (3). However, while Enserink reports scepticism regarding identification of a cause of Gulf War Illness after 10 years and \$155 million spent on research, the research on dioxins spanned decades and cost more than \$2 billion (4).

Lacking mention of these findings following the Vietnam War, Enserink's report appears overly pessimistic about the prospects for health research following the Gulf War. The concern is that we might lose opportunities to identify a new environmental exposure factor in order to to prevent recurrent problems, if we fail to continue use of present scientific tools.

GREG KARRAS

Communities for a Better Environment (CBE), 1611 Telegraph Avenue, Suite 450, Oakland, CA 94612, USA; e-mail: gkarras@cbecal.org

References and Notes

- 1. J. Kaiser, Science 290, 1071 (2000).
- G. Karras, in Persistent, Bioaccumulative, and Toxic Chemicals II, Assessment and New Chemicals, R.L. Lipnick et al., Eds. (American Chemical Society, Washington, DC, 2001).
- 3. J. Kaiser, M. Enserink, Science 290, 2053 (2000).
- D.R. Zook, C. Rappe, in *Dioxins and Health*, A. Schecter, ed. (Plenum, New York, 1994).

I WISH TO TAKE ISSUE WITH THE DESCRIPTION IN

Martin Enserink's New Focus article on Gulf War Illness (GWI) of our discovery (1, 2) of Mycoplasma fermentans in ~40% of GWI patients that "it is not clear whether M. fermentans really causes disease." In fact, there are numerous peer-reviewed papers on this issue (reviewed in 3), and a patent supported by the U.S. Army has been issued entitled "Pathogenic Mycoplasma" (4). Studies of its pathogenic properties have been published by the Armed Forces Institute of Pathology showing that healthy monkeys injected with M. fermentans developed a chronic illness that progresses to become fatal (5). Also, civilian patients with similar symptoms also show high rates of infection (2, 6, 7). M. fermentans fulfills almost all of the criteria of pathogenicity (8), including recovery on specific antibiotics (7). On the basis of this

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 6 months or issues of general interest. They can be submitted by e-mail (science_letters@aaas.org), the Web (www.letter2science.org), or regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

information the Department of Veterans Affairs established Cooperative Clinical Study Program #475, a blinded, placebo-controlled study on the effects of antibiotic treatment on GWI patients with *M. fermentans* infections. Although the clinical results of this study are not yet available, the laboratory entry criteria to the study indicate that a high percentage of GWI patients have systemic *M. fermentans* infections.

GARTH NICOLSON

The Institute for Molecular Medicine, 15162 Triton Lane, Huntington Beach, CA 92649, USA; e-mail: gnicolson@immed.org

References and Notes

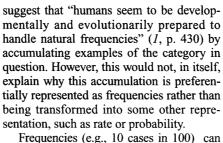
- G. L. Nicolson, N. L. Nicolson. *Intern. J. Occup. Med. Immunol. Tox.* 5, 69 (1996); G. L. Nicolson, N. L. Nicolson, M. Nasralla. *Intern. J. Med.* 1, 80 (1998).
- A. Vojdani, A. R. Franco, J. Chronic Fatigue Syndr. 5,187 (1999).
- G. L. Nicolson et al., J. Chronic Fatigue Syndr. 6 (3/4), 23 (2000).
- S.C. Lo. Pathogenic Mycoplasma. U.S. Patent 5,242,820. September 7, 1993.
- S. C. Lo, et al. Clin. Infect. Diseases 17 (S1), 283 (1993).
- M. Nasralla, J. Haier, G. L. Nicolson. Eur. J. Clin. Microbiol. Infect. Dis. 18, 859 (1999).
- M. Nasralla, J. Haier, G. L. Nicolson. *Intern. J. Med. Biol. Environ.* 28, 15 (2000).
- D. Taylor-Robinson. Clin. Infect. Diseases 23, 671 (1996).

Statistics: What Seems Natural?

WHICH STATISTICAL DATA SEEM EASIER TO understand, 10 cases in 100, or 10%? In their Policy Forum "Communicating statistical information" (Science's Compass, 22 Dec., p. 2261), U. Hoffrage and colleagues offer persuasive evidence that both experts and novices find it to be the former. When prevalence, sensitivity, and false positive rates are given as probabilities (e.g., 10%), most physicians misinterpret the information in a way that could be potentially disastrous for the patient, but when they are presented as "natural frequencies" (e.g., 10 cases in 100), the physicians' performance is dramatically better. The authors suggest ways to improve both communication of statistical information and medical education by using frequencies rather than probabilities.

The discussion by Hoffrage et al. leaves open the question as to why this is the case. Elsewhere, Gigerenzer and Hoffrage

SCIENCE'S COMPASS



Frequencies (e.g., 10 cases in 100) can be thought of as a subcollection (with a numerosity of 10) in a collection (with a numerosity of 100). I have suggested that we are born with a specialized capacity for representing collections and their numerosities (2). The evidence for this comes from a range of studies showing that infants, even in the first week of life, are sensitive to changes in the numerosity of a collection of visual objects (3) and that, at 6 months, they are able to form arithmetical expectations on the basis of adding an object to a collection or taking it away (4). The almost universal use of fingers as the representative collection in counting and arithmetic suggests that collections and numerosities form the basis of later representations also (2). This suggestion has been supported by recent brain-imaging evidence showing that key number areas are closely connected to the finger circuit in the intraparietal sulci (5).

Of course, the big developmental gap between the capacities of young children and the performance of adult decisionmakers is typically filled by an education system that teaches children about collec-

"[W]e are born
with a specialized
capacity for
representing
collections and
their
numerosities."

tions and numerosities far more than about probability. It is thus plausible that educational practices are, in part, responsible for the biases Hoffrage *et al.* report. However, there is indirect evidence that probability concepts are intrinsically difficult for humans. Although the computational techniques required by probabilities of the type described by the authors would have been available to the ancient Greeks, an understanding of the concepts began only with Girolamo Cardano's *Liber de ludo aleae*

(1525, published in 1663) and in the correspondence between Pascal and Fermat about games of chance in 1654.

BRIAN BUTTERWORTH

Institute of Cognitive Neuroscience, University College London, Alexandra House, 17 Queen Square, London WC1N 3AR, UK; e-mail: b.butterworth @ucl.ac.uk

References and Notes

- G. Gigerenzer, U. Hoffrage, Psychol. Rev. 106, 425 (1999).
- B. Butterworth, What Counts: How Every Brain Is Hardwired for Math (Free Press, New York, 1999) [in the UK, The Mathematical Brain (Macmillan, London, 1999)].
- S. E. Antell, D. P. Keating, Child Devel. 54, 695 (1983);
 P. Starkey, R. G. Cooper Jr., Science 210, 1033 (1980);
 E. Van Loosbroek, A. W. Smitsman, Devel. Psychol. 26, 916 (1990).
- 4. K. Wynn, Nature 358, 749 (1992).
- S. Dehaene, E. Spelke, P. Pinel, R. Stanescu, S. Tsivkin, Science 284, 970 (1999); B. Butterworth, Science 284, 928 (1999).

Response

BUTTERWORTH SUGGESTS THAT NATURAL

frequencies facilitate reasoning because people are born with a specialized capacity for processing collections of discrete objects, rather than probabilities and fractions. There is certainly evolutionary and developmental evidence supporting this explanation, but there is also a second, more specific reason: Bayesian computations are simpler when information is represented in natural frequencies than in probabilities, percentages, or relative frequencies (1, 2).

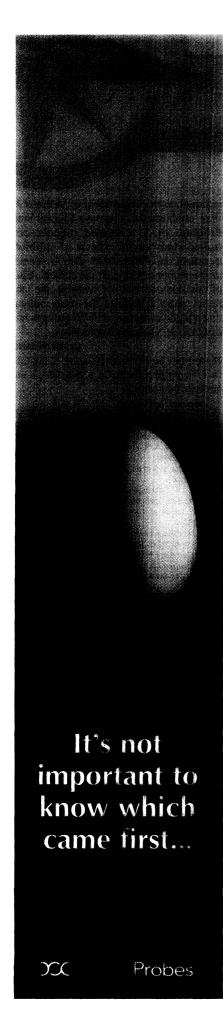
With natural frequencies, people can calculate the conditional probability of a hypothesis H (e.g., the occurrence of cancer) given data D (e.g., a positive test) simply:

$$p(H|D) = a/(a+b)$$
 (1)

where a is the natural frequency of people with cancer who tested positive and b is that of people without cancer who tested positive. In the colorectoral cancer example used in our Policy Forum, a equals 15 and b equals 300 people, respectively. Communicated this way, it is easy to see that 15 out of the 315 people who tested positive actually have cancer. In contrast, when the same information is communicated in terms of conditional probabilities, as is common practice, the calculation is complicated:

$$p(H|D) = p(H)p(D|H)/p(H)p(D|H) + p(\neg H)p(D|\neg H)$$
(2)

Equation 2 is known as Bayes' rule. In our example, applying this rule would require the nontrivial computation (0.003)(0.5)/[(0.03)(0.5) + (0.997)(0.03)]. The reason why natural frequencies facilitate Bayesian inference is because they retain in-



ULRICH HOFFRAGE, 1* SAMUEL LINDSEY, 2 RALPH HERTWIG, 1 GERD GIGERENZER1

¹Max Planck Institute for Human Development, Lentzeallee 94, Berlin 14195, Germany; ²Department of Psychology, University of Virginia, Charlottesville, VA 22903, USA

*To whom correspondence should be addressed. E-mail: hoffrage@mpib-berlin.mpg.de

References and Notes

- G. Gigerenzer, U. Hoffrage, Psychol. Rev. 102, 684 (1995).
- G. Gigerenzer, U. Hoffrage, Psychol. Rev. 106, 425 (1999).

The Bioenergetics of the Immune System

NUTRIENT CONSERVATION AT TIMES OF starvation is an important survival trait for any organism. This principle is elegantly demonstrated in the Report entitled "Survival for immunity: The price of immune system activation for bumblebee workers" by Yannick Moret and Paul Schmid-Hempel (10 Nov., p. 1166), in which the survival of bumblebees was significantly impaired by their inability to perform this function. Activation of the innate immune system of these insects caused increased mortality only when their nutrient intake was limited. The most likely cause of this increased mortality was the energy demands imposed by the immune response in the context of a paucity of energy substrate.

In mammals, it has been shown that the adipocyte-derived hormone leptin is a regulator of metabolism and bodyweight (1). An important physiological role of leptin is as a signal of starvation, in that a falling serum leptin concentration leads to neurohumoral and behavioral changes that seek to preserve limited energy reserves for immediately vital functions (2). It has been proposed that reduced leptin levels during conditions of starvation lead to impaired reproductive, thermogenic (2), and immune capabilities (3). One of the key features of the innate immune response is that its response is the same on each subsequent exposure to a certain stimulus, whereas the cognate immune system shows markedly different responses upon subsequent reexposure to a particular antigen. Indeed, the cognate immune response is far more energy-expensive than the in-



Bumblebees pay a hidden survival cost when their immune system is activated.

nate response, because of the necessity for the large-scale antigen-specific clonal expansion of lymphocytes.

The data presented by Moret and Schid-Hemple suggest that down-regulation of the immune response observed during starvation is an adaptive process, preserving vital energy supplies for cardiac and cerebral metabolism and hence the survival of the organism. We would propose that a falling leptin concentration during starvation causes a much-reduced cognate immune response. To date, leptin expression has been detected as far back in the evolutionary tree as ectotherms (4). It is interesting to note that it is at this stage of evolution that a cognate immune system emerged, with its intrinsic energy-expensive requirements.

GRAHAM M. LORD, 1* GIUSEPPE MATARESE, 2 JANE K. HOWARD, 3 ROBERT I. LECHLER 1

Departments of ¹Immunology and ³Endocrinology, Imperial College School of Medicine, Hammersmith, Du Cane Road, London W12 ONN, UK; ²Department of Biology and Cellular and Molecular Pathology, Cattedra di Immunologia, University of Naples "Federico II," via S. Pansini 5, Naples 80131, Italy.

*To whom correspondence should be addressed. E-mail: g.lord@ic.ac.uk

References and Notes

- 1. I. L. Halaas et al., Science 269, 543 (1995).
- 2. R. S. Ahima et al., Nature 382, 250 (1996)
- G. M. Lord et al., Nature 394, 897 (1998); J. K. Howard et al., J. Clin. Invest. 104, 1051 (1999).
- R. M. Johnson, T. M. Johnson, R. L. Londraville, J. Exp. Zool. 286, 718 (2000).

Response

SEVERAL OF THE THEORETICAL AND EMPIRICAL

points that Lord and colleagues touch on deserve further elucidation. For example, they refer to down-regulation of the immune defense under stressful conditions. This has not only been suggested for humans, performing excessive physical exercise (I), but has been experimentally demonstrated, for example, with birds forced to increase parental effort (2), or

