Volume 1 Number 1 1999

Of Significance...

A Topical Journal of The Association of Public Data Users

Statistical Literacy

APDU Association of Public Data Users

Of Significance...

Of Significance... represents an effort on the part of the Association of Public Data Users (APDU) to bring together a variety of perspectives on specific issues important to the collectors, disseminators and users of public data. Each issue will focus on a single theme related to the mission and interests of APDU members.

Information to Authors:
Of Significance... is published twice a year. Readers interested in submitting articles for possible publication should consult the APDU web site for specific issue themes and contact person. Inquiries regarding future issue themes and submission requirements can also be obtained by calling or writing APDU Administrative Headquarters.

Price:
Non-U.S. individuals and organizations:
Yearly subscription: $80 (U.S.)
Single issues: $45.00 (U.S.)
U.S. individuals and organizations receive OF SIGNIFICANCE as a benefit of membership. In addition to *Of Significance...*, member benefits include a quarterly newsletter, conference discounts, and occasional special publications. Information on member categories and costs can be obtained by contacting APDU Administrative Headquarters or checking the APDU web site.

Back Issues and Claims:
Information about availability of issues may be obtained from APDU Administrative Headquarters. Claims for undelivered copies must be made no later than six months after publication. The publisher will supply missing copies when losses have been sustained in transit and when the reserve stock will permit.

Name of publication: Of Significance...
Issue: Volume 1, Number 1, 1999
Frequency: Two issues per year, June and December

APDU Administrative Headquarters
Teresa Hall Allen, Chief Administrator
P.O. Box 12538
Arlington, VA 22219
Phone: (703) 807-2327
FAX: (703) 528-2857
E-mail: apdu@smdi.com
URL: www.apdu.org

Guest Editor for Vol. 1:1: Wendy Treadwell
General Editor: Deirdre Gaquin
Production Editor: Janet M. Roseen

Contents

How to Help the Public (and Yourselves) Know the Truth

VICTOR COHN

Abstract: There is wide interest today among both journalists and scientists in helping the public learn what to believe, in helping the public learn to ask questions, in helping the public navigate through the shoals of today's reporting about a thousand subjects. For data professionals the questions are the same. What can you believe and disseminate, and how can you tell the difference? How can you help the public learn what to believe?

Adapted from Victor Cohn, *News & Numbers: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields* (Iowa State University Press, ©1989, rev. 1994.)

Everybody has data to make us believe something or other. Or at least they have some numbers. Well, not everybody. A lot of folks say *such and such* is true, without giving us any data or numbers at all. For journalists like me and a public barraged with ever-changing news about coffee, tea, breast cancer, fats in the diet, global warming, you name it, the question is: "What can we believe?"

For data professionals, the question is the same. What can you believe and disseminate, and how can you tell the difference? Plus another question: how can you help the public learn what to believe? How can you help folks separate the probable or possible truth from the probable trash in all they see, read, and hear?

There is a lot of trash. And there is a lot of hype, sometimes quite unintended, as TV, radio, and print media battle for the public's attention. Some years ago, tongue only partly in cheek, I said that there are only two kinds of health news: New Hope and No Hope. New Hope and No Hope get on page one of the evening news. The stuff in the middle, actually the bulk of health news, too often gets ignored or buried. Some examples:

- A *Washington Post* headline on an Associated Press story on research using monkeys once read, "Technique May Aid Acceptance of Transplants." (*Post*, 8/6/97) The same story was trumpeted on TV news programs. Down in the third paragraph of the AP story, the investigator did caution "that it will take more years of research before the technique can be used in humans." But even more striking, down in the very last paragraph we learned that the technique worked—at least for 150 days, we aren't told more—in only two of eight monkeys in which it was tried. This is not what any scientist would call good evidence.
- The *New York Times* once told us, "New Medicine May Help to Cut Sting of Common Colds in Half" (*Times*, 10/1/97). The *ABC Evening News* had told us the same thing the evening before, with few reservations (ABC, 9/30/97). Good news? Well, it turns out that the new medicine was sprayed into volunteers' noses either seven hours before or twelve hours after exposure to a cold virus. So how could this spray prevent a cold if we must take it just before or after an exposure we don't know about?
- A *Newsweek* cover asked, "The End of AIDS?" (*Newsweek*, 12/2/96). Okay, it added "Not Yet," but the overall impression was another new hope. Yes, there had been a dramatic stride against AIDS. Some patients had at least staved off the disease's ravages by a daunting and expensive array of drugs taken several times daily. But the drugs weren't working for everyone, they were already failing in many patients, and they were so costly that many couldn't get them. Listen to Mary Fisher, a woman—widely seen on TV a few years back—who contracted AIDS from a blood transfusion (*Post*, 6/9/97). She said the drugs were making her so ill that she was struggling to care whether she lived or died. She said: "I'm tired. I can't sleep. The

Victor Cohn, former science editor of The Washington Post, is the author of NEWS & NUMBERS: A Guide to Reporting Statistical Claims and Controversies in Health and Other Fields.

treatment is worse than the disease." That didn't make a news magazine cover.

- When National Cancer Institute doctors in 1985 treated their first patients with Interleukin-2—a substance intended to turn white blood cells into anti-cancer agents—typical headlines read "Killer Cells," "Highly Promising," and the like. The impression was that a new era was nigh. But promising has remained the best that can be said of this compound, with few tumor disappearances reported.
- The mid-'80s saw a wave of implants of so-called artificial hearts, dominating the medical news, though none could pump blood like a real heart, and all were connected to bulky bedside machinery. None worked, if the test of what works is life.
- Kidney and heart transplants, kidney dialysis, heart pacemakers—all, we read or hear, are extending lives. Indeed they are. But—and we don't hear much about this—all often fail. When I was writing for the *Washington Post,* a Maryland woman once phoned me to say, "My son had to have his first pacemaker at age 25. Now he's wearing his fourth. They were infected! The miracles aren't always miracles."
- Anchorman Peter Jennings told us one night on the *ABC Evening News* of a newly emerging computer-generated medical diagnosis, where a doctor puts your symptoms into a computer and gets back an answer (ABC, 10/6/97). Jennings said, "It is fast, it is flawless." Nothing in medicine is flawless.

The overall lesson is that medical *miracles* and *wonders* are rare, and a lot of stuff is reported that isn't quite ready for prime time, or gets ballyhooed beyond its current worth. One night on *CNN Headline News,* I heard it said of the parents of a very sick child: "These parents wait for medical science to work its wonders" (CNN, 10/16/96). Yet—unless a proven treatment were available, which it was not—these parents were doomed to disappointment. And disappointment and distrust of mainstream medicine is the attitude of millions of Americans today as they have too often been fed *New Hopes* that haven't panned out.

I am not saying that all reporting on health issues is bad. Indeed, much of it is very good, and the public today is better informed than ever about health and medicine. But we all need to learn to be better judges of what to believe, and, most important, what to act on for our own benefit.

How can we do this when we're not scientists, we're not doctors, we're not experts on all these things? Well, we can. We can apply some of the tests that scientists apply when they judge each other's reports.

To be specific, we can go a long way toward deciding what to believe, a long way toward discerning probable facts from probable junk, a long way toward judging claims and even statistics that are thrown at us by applying some basic ideas that apply to all science, all medical studies, and virtually all knowledge of the universe.

The first thing we have to understand is the ***certainty of uncertainty*** about virtually everything in science and medicine. The fact is that all science is almost always uncertain, or uncertain to a degree. Nature is complex. People, part of nature, are worse. Research is difficult, observation is inexact, all studies have some flaws. So science is always an evolving story. Almost all anyone can say about the behavior of atoms or people is that there is a strong probability that such-and-such is true, and we may know more tomorrow.

Dr. Arnold Relman, then editor of the *New England Journal of Medicine,* said, "Almost anything we say is the best we can do, our best opinion at the moment, and things may be updated in the future" (Relman).

This tells us why we so often hear people say, "Why do *they* say one thing today and another tomorrow." It tells us why things so often seem settled one way today and another later, and why so much is debated, whether the effects of global warming or a high fat diet or almost any medical treatment, and why so much is in dispute, whether coffee is or is not harmful, nuclear power is or is not a danger. *Et cetera. Et cetera.*

Why so much uncertainty? Many reasons. The sheer difficulty of much research. Lack of funds to do enough research, especially long, continuing observations of large populations. The ethical obstacles to using human beings as guinea pigs. But also progress. Medical science is improving, and old ideas are constantly being put to the test.

Take breast cancer treatment. For nearly one hundred years, surgeons believed that only a radical mastectomy—complete amputation of the breast and nearby glands—was the only treatment for breast cancer. Only in the 1970s did clinical trials begin to show that less drastic treatments can often be equally effective, and, by now, often better.

A few decades ago Alzheimer's disease was considered to be a rare condition. Now it is recognized as the main cause of conditions once ascribed to senility or arteriosclerosis, though arteriosclerosis and strokes can still have a similar result and it is not always easy to tell which is which. Some of you may remember ulcer diets and strict bed rest. Long lists can be made of treatments that once seemed right, only to be proved wrong by today's more rigorous comparisons of one treatment with another.

So uncertainty most often prevails, and we must live with it. But can we? When personal action is required, when we have to decide what kind of advice to take about our own health—or, as a society, when public health is involved—there is a way we can still use all the information available to make sensible decisions.

Scientists deal with uncertainty by invoking ***probability***—and forgive me now if I repeat in elementary form some of the statistical principles that many of you know so well, but others may have forgotten. Scientists measure probability by some reliable formulas. To ***be statistically significant***, and not just the result of pure chance, the same result must appear again and again. When it does, that's ***reliability***.

The laws of probability and chance also tell us to expect some unusual, even impossible sounding, events. Like going on a trip to Timbuktu and meeting our next door neighbor. Or tossing heads or tails several times in a row. These things happen. Similarly, nature will randomly produce many alarming clusters of cancers or birth defects—in a block or school or neighborhood—which actually have no cause but nature's coin tossing. This is not to say that such clusters may not sometimes have a real cause. But a cluster alone does not constitute evidence.

There is something else to remember when someone says, "How do they know this stuff isn't causing harm?" Science cannot prove a negative. No one can prove that Little Green Men from Mars have not visited earth. The burden of proof, the necessity to produce real physical evidence, to produce real, reliable numbers should be on those who say something is true, whether visits of Little Green Men or the worth of some substance that some doctor says is curing cancers no one else can cure.

Be wary when you read or hear of studies with only a small number of cases. ***Power*** in science is the power of large numbers. The greater the number of cases or subjects studied, the greater a conclusion's power and probable truth.

Dartmouth doctors in 1984 placed catheters into the skulls of four Alzheimer's patients and injected a theoretically promising drug. According to the patient's families—not independent observers—three showed improvements and the fourth held his own. The university unwisely held a news conference, the story was widely trumpeted, and within two months the medical center logged 2,600 phone calls, mainly from desperate families. Further study showed that the drug didn't work (Winsten). Four patients don't tell us much. Even forty might not. Often far larger numbers are needed.

If a doctor tells you, "I've been using this drug for ten years and never had a bad result," remember this: the likelihood that a 30-to-39-year-old woman will suffer a heart attack while taking birth control pills was once described as one per 18,000 women per year. (Today, with better pills, it may be even less.) But—at one per 18,000—to be 95% sure of seeing this even once in a single year, a physician would have to observe 54,000 women (Bessler).

This tells us why we so often learn of a drug's harmful side effects only after it has been studied and approved and is being used by many thousands. This also tells us why so many medical scientists shudder when they keep seeing Congressional pressure—reflected in legislation—to force the Food and Drug Administration to speed up the drug approval process rather than do slow, calculated testing in as many persons as possible.

Of course we all want new, possibly helpful, drugs on the market as soon as possible. And sometimes we decide we must take a chance that we will do more good than harm. But it is often a difficult decision, for, surely, some new drugs will do unexpected harm.

Scientists judging the result of a study also look for absence of ***bias.*** Bias in science is a little different from human bias, which also exists. Bias in this sense means seeing false associations and reaching unreliable conclusions by failing to consider some important influences, like the effects of age, gender, occupation, nationality, race, incomes, health, or various behaviors.

Like smoking. Among women suffering ill effects from birth control pills, older age and smoking were for years the two most ignored variables—conditions—that helped promote or cause these effects.

When someone says such-and-such is caused by so-and-so, we need to ask, "Did you ignore any other possible causes?"

Another common pitfall is ignoring the effects of ***variation*** or ***variability.*** Everything that is measured or studied varies from measurement to measurement. No two studies of exactly the same thing have exactly the same results, and often the results vary widely. Our physiologies fluctuate from minute to minute. People are different from each other. And people in different parts of the country may vary greatly. This is another reason why we can almost never rely on just one study to tell us anything. We want to see a study repeated and repeated before putting down our chips.

And ***studies*** vary in reliability from the least to the generally most believable. Someone says, "I've done a study." But all studies are not equal. They vary in worth—from simple anecdotes or stories about this patient or that, interesting but not evidence, going on to more systematic observation or eyeballing, sometimes all that is possible but still not very reliable. And finally proceeding to true experiments, especially those meeting the gold standard—comparing one population or sample or group of patients with another under different treatments or different conditions.

What does all this tell us? It tells us that we have to pay attention. It tells us that studies typically vary widely in results. This is expected, for there are commonly conflicting studies on many subjects. The most believable studies and observations are those repeated and repeated among different populations with much the same result, and supported, if possible, by animal or other biologic evidence.

It tells us that we should beware of studies that have only a handful of cases. It tells us that proof of anything requires more than anecdotes or miracle stories, however dramatic. It tells us not to rely on any one report, but to look for a consensus, an agreement among the best studies and the best observers, and to look for reports that describe not only today's finding but that put it in the context of other or past reports and tell us whether or not there is general agreement. Or, still, wide disagreement.

It tells us to ask questions. The most important question to ask: how do you, or they, know? Are you just telling us something you've "observed" or "found to be true?" Or have there been some good studies or experiments? Were they conducted on a lot of people or just a few? What are the data, the numbers? How sure are the results, how valid or accurate? How reliable or reproducible are they? Have results been fairly consistent from study to study? Do people agree about this? Who disagrees? And why?

To sum up, I have been emphasizing your role as data collectors and disseminators in deciding what can be believed and what is worth communicating to the public. But reporting responsibly should also be the responsibility of researchers and scientists.

Researchers once wrote in the *New England Journal of Medicine* that "one to three drinks" a day may help protect against heart attacks. They defined a drink as 13.2 grams of alcohol on average (Gaziano). But with whiskey, wine, and beer all at various proofs or alcohol content, neither the NEJM article nor any news reports I read or saw told how much daily booze, wine, or beer one should consume to drink no more than the prescribed grams. A freely pouring drinker could pour far more.

I believe that medical editors should recognize that ordinary people are now part of their audience via scores of news reporters in print and on TV. So are physicians who treat people. Do you think all the physicians who read that article were immediately able to convert the requisite number of grams into ready advice for their patients?

I believe there is actually wide interest today among both journalists and scientists in helping the public learn what to believe, in helping the public learn to ask questions, in helping the public navigate through the shoals of today's reporting about a thousand subjects. This very much includes all the hype and hyperbole we're hearing and reading about alternative and natural supplements and medicines. There are surely gems among the hype—and hopefully studies now under way will help reveal them—but there are also dangers.

Here and on every subject, help us, the journalists, help us, the public, learn to live with all we read and hear and all we must cope with.

References

ABC Evening News, October 6, 1997.
ABC Evening News, September 30, 1997.
"Activist Calls AIDS Treatment Worse Than the Disease," *Washington Post,* June 9, 1997.
Bessler, Stuart A., Ph.D. (Syntex), personal communication, November 28, 1984.
CNN Headlines News, October 16, 1996.
"New Medicine May Help to Cut Sting of Common Colds in Half," *New York Times,* October 1, 1997.
Newsweek, December 2,1996.
Relman, Arnold, M.D., personal communication, November 30, 1984.
"Technique May Aid Acceptance of Transplants," *Washington Post,* August 6, 1997.
Winsten, Jay A., Ph.D., "Science and the Media: the Boundaries of Truth," *Health Affairs,* Spring 1985.
Gaziano, J. Michael *et al*., "Moderate Alcohol Intake, Increased Levels of High Density Lipoprotein and Its Subfractions and Decreased Risk of Myocardial Infarction. *New England Journal of Medicine,* December 16, 1993.