"Placebo Effect is More Myth than Science, Study Says"

by Gina Kolata, The New York Times, May 24, 2001, p. A20.

"Is the Placebo Powerless? An Analysis of Clinical Trials Comparing Placebo with No Treatment"

by Asbjorn Hrobjartsson and Peter C. Gotzsche, New England Journal of Medicine, 344 (May 24, 2001), p. 1594-1602.

"The Powerful Placebo and the Wizard of Oz"

by John C. Bailar III, New England Journal of Medicine, 344, p. 1594-1602.

Research reported in the New England Journal of Medicine points to a startling conclusion: the well-known placebo effect may be a statistical myth! Dr. Hrobjartsson is quoted in the Times as saying that no previous published research had distinguished between the placebo effect and the natural variations in symptoms of a disease. He and Dr. Gotzsche therefore undertook an extensive literature review to find studies that included both a placebo group and an untreated group. If the placebo effect were real, they reasoned, then it ought to show up in comparisons of these two groups. Ultimately, they turned up 114 suitable studies, involving a total of 7500 patients and 40 different medical conditions. When the results were tallied, patients on placebo treatment were not found to do better than those left untreated.

The NEJM article gives more details on the analysis. The researchers first classified the studies. Some involved binary responses (the patient's condition either improved or did not), while others had continuous responses. In some, the response was based on the patients' subjective perceptions, whereas others measured physiological responses. Also, three types of placebos were distinguished:
pharmacological (e.g., pills), physical (e.g., manipulation), and psychological (e.g., a conversation).

For binary outcomes, there was no significant difference between placebo and no treatment. Considering subjective and objective responses separately gave results similar to the overall results. On the other hand, for continuous outcomes there was a significant difference overall between placebo and no treatment groups. However, when the continuous results were broken down by response type, the subjective responses showed a significant difference, while the objective responses did not. Furthermore, it was found that the effect decreased as the size of the trials increased. The researchers suggested a possible bias in the smaller trials; the patients' subjective reports of improvement in smaller studies may reflect a desire to please their doctors. Patterns similar to the above were observed when pharmacological, physical and psychological placebos were considered separately.

Hrobjartsson and Gotzsche conclude that placebo treatment has no use beyond clinical trials experiments. John Bailar's editorial accompanying the research article compares the power of the placebo to the power of the Wizard of Oz, which depended on no one looking behind the curtain. Still, Bailar seemed uncomfortable with a complete rejection of placebos. He holds out some hope that they may still be useful in specific settings, such as pain relief.

The Times coverage includes interesting reactions from several other experts. Berkeley statistician David Freedman pointed out that pooling data from many studies in a meta-analysis can sometimes produce misleading results. He is quoted as saying: "I just don't find this report to be very persuasive. The evidence of a placebo effect is maybe a bit less than I thought it was, but I think there is a big effect in many circumstances." On the other hand Dr. Donald Berry, a statistician at the M. D. Anderson Cancer Center in Houston, said of "I believe it [the new finding]. In fact, I have long believed that the placebo effect is nothing more than a regression effect."

For a touch of dramatic irony, we must mention the article "Barry Bonds and the Placebo Effect," which Princeton mathematician Jordan Ellenberg wrote for the his "Do the Math" feature in the July 12, 2001 edition of the electronic news magazine Slate.

slate.msn.com/?id=111848

Ellenberg presents a nice discussion of regression to the mean, including some quotes from Francis Galton's memoirs. At the time of Ellenberg's article, Barry Bonds had just set a record for most home runs hit before the All-Star break (39), and he was on pace to hit 72. "But," Ellenberg observed, "Barry Bonds isn't going to hit 72 home runs for the same reason that there might be no such thing as the placebo effect."

"How Colleges Reject the Top Applicants -- and Boost their Status"

by Daniel Golden, The Wall Street Journal, May 29, 2001, p. A1.

Critics of college rankings often express the concern that schools may skew their policies in response to the formulas used to rank them. Indeed, according to this article, some colleges have begun to place applicants perceived as "overqualified" on the waiting list, while offering admission to students who are objectively less qualified but considered more likely to enroll. The practice has become increasingly common among schools perceived as just below the top tier. The article says that last year Franklin and Marshall College rejected 140 of its top applicants, because they had not interviewed with the school or
otherwise demonstrated interest in attending. Past admissions experience suggested that such students were not likely to enroll.

It may be that such schools are just being realistic. But there is a curious benefit to such policies, which have the effect of increasing acceptance rate and admissions "yield" (the percentage of those admitted who ultimately enroll). The figures count for 1/4 of the selectivity score in the popular *U.S. News & World Report* College Rankings. The article estimates that changes in these numbers could move a school up or down several positions in the rankings. You can find more details about the *U.S. News* formula at their Web site

www.usnews.com/usnews/edu/college/rankings/about/weight.htm

The schools have additional data to back up their policies. Many now employ consulting firms to help manage the admissions process. The consultants have developed statistical models that use intended major, extracurricular activities, and other demographic variables to predict whether an applicant will enroll if accepted. In some of these models, when an applicant's test scores exceed the median for the school, the predicted chance of enrolling goes down.

"Connoisseurs of Chaos Offer A Valuable Product: Randomness"

by George Johnson, *The New York Times*, June 12, 2001, p. F1.

John von Neumann once remarked that "Anyone who considers arithmetic methods of producing random digits is, of course, in a state of sin." Nevertheless, computer scientists have over the years developed many algorithms to approximate randomness with so-called pseudorandom number generators. The article summarizes the history of attempts to produce random digits, including the RAND Corporation's famous 1955 publication "One Million Random Digits with 100,000 Normal Deviates," and von Neumann's own "middle square" algorithm.

Although pseudo-random numbers have in the past proved sufficient for simulation and sampling experiments, the article notes that modern cryptological applications require something more like true randomness. It references three Web sites that employ physical mechanisms rather than formal algorithms to provide random digits.

The Hotbits project

www.fourmilab.ch/hotbits

generates random digits by using radioactive decay. If you connect to this site, you can listen to the Geiger counter ticks.

The Random.org site

www.random.org

samples atmospheric noise by using a radio tuned between stations. This sounds closest to the original RAND project, which sampled the output of an electronic pulse source. The site provides a general discussion of random numbers and many links to further information.
The Lavarand site

lavarand.sgi.com

is based on the idea that a lava lamp is a chaotic system; that is, it exhibits "sensitive dependence on initial conditions." Their technique employs six different colored six lamps, which are photographed with a digital camera at fixed time intervals. The resulting pixel map is converted to bits. The *Times* article features a picture of a lava lamp!

"Victim Poll on Violent Crime Finds 15% Drop Last Year"

by Fox Butterfield, *The New York Times*, June 14, 2001, A16.

This article describes two apparently conflicting reports on trends in violent crime. The Justice Department's newly released National Crime Victimization Survey (NCVS) found a 15% drop in violent crime last year, the largest single-year drop since the survey was begun in 1973. On the other hand, in late May the FBI released its Uniform Crime Report (UCR), which indicated that violent crime held steady last year, after 8 consecutive years of decline (see: "U.S. crime figures were stable in 2000 after 8-year drop," *The New York Times*, May 31, 2001).

Because the UCR and NCVS had mirrored each other during previous years of decline, it may be that one of them has somehow gone awry this year. However, the article points out that the discrepancy may be attributable to differences in measurement methodology. First of all, the NCVS is based on interviews with crime victims, while the UCR is compiled from reports from law enforcement groups. Also, the definitions of "violent crime" differ somewhat. The NCVS covers rape, sexual assault, robbery, aggravated assault and simple assault. It does not include murder, which is covered by the UCR. On the other hand, the UCR does not include simple assault.

Of course, the Justice Department is fully aware of these differences. On their Web site

www.ojp.usdoj.gov/bjs/abstract/ntmc.htm

you can read an explanation for different aspects of crime reflected in the NCVS and UCR, and how the two measures actually complement each other. One advantage cited for the NCVS is that it includes crimes that are not reported to law enforcement agencies. On the other hand, since it is based on survey data, it is subject to sampling error.

Another explanation for the discrepancy was proposed by Iain Murray of the Statistical Assessment Service (STATS) in his article "Good News! More People are Reporting Crimes," (*The Washington Post*, July 15, 2001, p. B3). Murray hypothesizes that crime has now declined to the point where it is no longer accepted as the norm, and people have begun reporting crimes that might have gone unreported when criminal activity was more widespread. You can find the full text of his article on the STATS Web site.

www.stats.org/statswork/crime-washpost.htm
This summer saw a dramatic controversy in major league baseball. Sandy Alderson, baseball's chief of operations, was worried that games were running to long. In what became a problematic phrase, he instructed umpires to "hunt for strikes" in order to decrease the number of pitches that had to be thrown. Alderson cited a two-month study that purportedly found a "correlation between very high pitch counts and a misapplication of the strike zone." Pitch counts were averaging in the 280s, and he felt that figures in the 270s were attainable through more accurate umpiring. He said "I've told those averaging around 310, that's unacceptable and it's evidence of a very small strike zone and we have videotape to support it."

Needless to say, the umpires didn't appreciate the criticism. They had been told early in the season to follow more strictly the description of strike zone given in the official rule book, in particular calling strikes on higher pitches and on pitches on the inside corner of the plate. But the umpires now felt they were being told to twist the rules in order to finish games on time. The umpire's union filed a grievance, and the commissioner's office soon capitulated, issuing a statement that it would not use pitch counts and averages to judge an umpire's performance. The union then withdrew its grievance, and controversy was suddenly over -- for now, in any case.

Of course, we are talking about America's pastime here, so you can find many more commentaries and op-ed pieces in the news from this period. Here is a link to some data and graphics related to pitch counts, compiled by the Baseball Prospectus Online:

www.baseballprospectus.com/news/20010719aim.html

The summer was also the "Summer of the Shark," as *Time* magazine announced on its July 30 cover. Several highly publicized incidents and two deaths fueled a popular perception that shark attacks were increasing dramatically. But George Burgess, who maintains the International Shark Attack File (ISAF) at the Florida Museum of Natural History, tells us that there is no cause for panic. The ISAF Web site has the data to back him up:
There were 84 shark attacks in 2000 and so far there have been 52 in 2001. This year's total may actually turn out to be lower than last year's. More than half of this year's attacks, have occurred in Florida, but as the Times explains, experts say this simply reflects an increase in beachgoers visiting the state. The ISAF Web site presents graphics that show the numbers of shark attacks roughly tracking population figures. Here is a link to the graph for Florida:

www.flmnh.ufl.edu/fish/Sharks/Statistics/flpop.htm

The ISAF also presents data to put the relative risk of shark attack in perspective. The figures on lightning strikes and accidents with tools may not seem compelling, but the comparison with alligator attacks certainly seems relevant. We learn that between 1948 to 1995, there were 218 alligator attacks on humans in Florida, 7 of which were fatal; by comparison, there were 276 shark attacks during that period, 6 of which were fatal. Even though these are comparable figures, the shark attacks loom larger in the public imagination.

In his op-ed piece, mathematician John Allen Paulos remarks that simple arithmetic shows that the risk of being attacked by a shark is very small. He cites several other risks that are greatly feared despite being very small, such as contracting the West Nile virus or being abducted by a stranger. He observes that the risks attributable to alcohol, for example, are much greater. Another popularly overestimated risk is that of "air-rage," for which Paulos provides some explicit calculations. There are about 3,500 incidents a year involving abusive passengers, only 10% of which are serious enough to warrant the airline taking action against the perpetrator. Paulos writes:

This may still sound like a dreadful problem. But close to 2 million Americans fly every day. That's about 700 million passengers annually. Dividing 700 million 3,500 incidents, or by 350 serious outbursts, we find that about one in 200,000 passengers is involved in any air rage incident annually, and only one in 2 million is involved in a serious one. Compare those figures to the behavior and arrest rates at sporting events, and you'll appreciate what a docile bunch air travelers are.

Paulos says the news media are often at fault for reporting risks in an overly dramatic way. For example, in the case of the recent wild fires in the western US, the area burned was reported in acres. He points out that 2 million acres sounds worse than 3100 square miles -- even though they represent the same area.

"True, False, Whatever; Physicians are Putting a Stop to the Publication of Misleading Drug Data"

by Stacy Schultz, *U.S. News & World Report*, September 17, 2001, p. 72-73.

"Scholarship, Authorship, and Accountability"

by Frank Davidoff MD, et al., *Journal of the American Medical Association*, 286, September 12, 2001, p. 1232-1234.

In her book *Tainted Truth: The Manipulation of Fact in America* (Simon & Schuster, 1994), Cynthia
Crossen showed that dangerous conflicts of interest can arise when corporations get involved in scientific research. The *U.S. News* article reports that major medical journals are now taking strong editorial stands against the influence that pharmaceutical companies exert over research and publication.

According to the article, 70 percent of the clinical drug trials conducted last year were sponsored by pharmaceutical companies. An example of what can go wrong is provided by the case of Celebrex, a drug now widely used to treat arthritis. Its popularity stems in part from research reported in the *Journal of the American Medical Association*. Results from a 6-month trial showed that the drug had fewer side effects than its competitors. But as *JAMA* editor Catherine DeAngelis later learned, the company actually had 12 months worth of data, and the full data set did not support the findings that were published.

Withholding data from investigators is not the only kind of abuse identified here. In other cases, companies designed experiments without input from clinicians, lobbied for favorable interpretation of the resulting data, and then tried to block publication of findings that were unfavorable to their products. Of course, investigators are dependent on funding, so the pressure to go along with a sponsor's wishes can be hard to ignore completely. But of this impedes scientific progress.

A joint editorial published this week in 12 major medical journals condemns such practices (the *JAMA* citation is given above). The editors state that they "will not review or publish articles based on conditions that allow the sponsor to have sole control over the data or to withhold publication ... [Research] contracts should give the researchers a substantial say in trial design, access to the raw data, responsibility for data analysis and interpretation, and the right to publish ..."