

Spinning Heads and Spinning News: How a Lack of Statistical Proficiency Affects Media Coverage

Rebecca Goldin¹

STATS and Department of Mathematics
George Mason University, Fairfax VA 22030

Abstract

News increasingly depends on a careful dissection of numbers. Statistics are everywhere, from how many people lack health insurance to how to improve math education. Yet for being so prevalent, statistics are poorly understood by the general public. In this essay, I'll illustrate how the press can misuse and even abuse statistics using examples of news coverage. Since news sources are the main avenue by which the public understands many public health issues, these misguided representations of science can actually shape public policy, legislation, and individual choices. It is essential that media writers understand basic concepts from statistics, epidemiology and more generally the methods of scientific discovery. The press is most powerful when it goes beyond politics and morality to point out what science says, what it doesn't, and what it can't.

Key Words: Statistical education, statistical literacy, journalism

1. Introduction

Reporting the news accurately has increasingly come to depend on a careful dissection of numbers, and the press often isn't up to the job. Statistics are everywhere — from how many people lack health insurance to how to improve math education — yet they are poorly understood by the general public and the press alike. As a result, the press sometimes misuses, or even abuses, statistics in its news coverage, and readers rarely notice. Since news sources are the main avenue by which the public understands many public health issues, these misguided representations of science can actually shape public policy, legislation, and individual choices.

In order to present issues of risk, data, or science accurately, it is essential that media writers understand basic statistical and epidemiological principles, as well as the methods of scientific discovery. The press is most powerful when it goes beyond politics and morality to point out what science says, what it doesn't, and what it can't.

What are the most common mistakes in media and press coverage? They can be as simple as computational errors, but the more damaging misrepresentations come from misunderstanding:

- The difference between causation and correlation;
- The meaning of *statistically significant*;
- Orders of magnitude, and the “prevalence” of a problem;
- Confounding factors;

¹ The author was partially supported by NSF-DMS grant #0606869 and the Keck Statistical Literacy Project

- Relative risk versus absolute risk;
- Margin of error; and
- The importance of scientific consensus.

Many of the examples below came from my observations as Director of Research at Statistical Assessment Service (STATS), a nonprofit affiliated with George Mason University whose goal is both to offer educational resources and to call attention to poor or exceptionally good media coverage. You can read many media critiques as well as get access to resources for journalists at www.stats.org.

2. Sourcing the Problem

There are profound cultural differences between the world of scientists and the world of journalists. At first glance, these differences don't have anything to do with statistics; in a few moments, I will show why it really does. First however, let me discuss the culture divide.

For a scientist, a new result is suspect. Why is this result different what has been commonly accepted in the literature? For a journalist, a new result is just that: news. It is something to let people know about. A scientist wants to check results by repeating experiments, establishing corroborating evidence, providing a mechanistic explanation, and having his or her work peer-reviewed. A journalist is under tremendous time pressure, with little incentive to check that the scientific findings pass the test for scientific rigor. If the science confirms the consensus, it isn't, at least not typically, news.

The immediate consequence of these different points of view is that scientists and journalists often communicate ineffectively. This miscommunication goes both ways, but I want to focus on the problems that journalists have in understanding scientists.

When there is a divide among scientists, as happens when a small group disagrees with the consensus, the media may represent two scientific views equally, as if they had equal weight within the scientific community, when in fact, they do not. In this way, media (and therefore mainstream opinions) can be manipulated by other interests, such as commercial interests, that want to suggest doubt when there is little in the scientific community (this tactic was used by the cigarette industry for many years), or by special interest groups such as environmental groups, whose interests benefit when the public believes over-exaggerated claims of risk.

The resulting fear of small risks can be as damaging to our society as our ignorance in the face of great risks. To give an example, while we spend our public resources (and airwaves) worrying that our children could be kidnapped, we are somewhat cavalier regarding the risk of driving. According to the Department of Justice in 2002, there were 115 children (under age 18) in 1999 who were victims of "stereotypical" kidnappings in the United States [Department of Justice, 2002]. In contrast, according to the Centers for Disease Control, in 2006 alone, 1,694 children ages 1-14 were killed thanks to motor vehicles, and the number is a whopping 6,508 if children 15-19 are also included.

Superficially, the difference in perspective between journalists and scientists seems only reflective of the nature of journalism compared to the nature of science itself. However, there is something deeper: the lack of understanding by journalists of basic statistical tenets leads journalists away from the right questions and into the world of cherry-picked

stories. Just as one datum does not make for statistics, one story or one study does not represent the whole. However, journalists often make precisely this mistake. A dramatic account of how one boy committed suicide while on an anti-depressant is far more appealing to an editor's eyes (and more likely to capture readers' interest) than a carefully laid out account of scientific studies on anti-depressant drugs and their effects on teens. A few rogue scientists with fringe ideas and one or two scientific studies can cast doubt on established scientific consensus based on hundreds of studies.

One of the main focal points of STATS is to improve how risk is discussed in the media. In this context, STATS paired with the Center for Health and Risk Communication at George Mason University to poll toxicologists about their perceptions of chemical risk and its coverage in the press. [STATS, 2009]. We leave aside a discussion of how these toxicologists ranked those chemicals that are making headlines in terms of risk, and turn to the toxicologists' opinions about the quality of media coverage. In this survey, which included government, industry, and university toxicologists, 96 percent of toxicologists say the media does a poor job distinguishing between correlation and causation, 97 percent say that the media does a poor job distinguishing between good and bad studies, and 95 percent say that the media does a poor job of explaining the risk/benefit trade off. In addition, 74 percent say the media gives too much weight to individual studies relative to overall evidence. Finally, the source of news coverage with the highest accuracy rating given by those toxicologists who had opinions about the matter was WebMD, followed by Wikipedia. Notably, it was not national newspapers, or even NPR or PBS.

While toxicologists do not necessarily represent all scientists, this poll is sobering: if the media wants to retain relevance, it must communicate scientifically, and in order to do that, it must adopt a scientific way of thinking. Statistics and statistical thinking form one of the basic building blocks for scientific reasoning as a whole.

3. Causation vs. Correlation

The difference between causation, when one thing causes another, and correlation, when two occurrences go hand-in-hand, is as basic as arithmetic to statisticians. Not so for journalists. At times, their conflation of these is transparent. Just a few months ago, for example, *The New York Times* reported that “lengthy television viewing in adolescence may raise the risk for depression in young adults.” [February 10, 2009] *US News and World Report* made a similar claim. Note the causal implication with the word “raise.” Keep in mind the word “may” in common parlance is not the same as in a technical context; the headline implies that the reported study offers evidence that television raises the risk of depression.

The news reports were based on a research study which found a correlation between television and subsequent depression: six percent of children who had watched less than three hours a day subsequently had “depressive symptoms” while 17 percent of those who had watched more than nine hours a day developed such symptoms.

But which children are watching nine hours of television a day? Are depressed kids more likely to sit in front of the screen than other kids? The question is simply not posed, not even by first tier national news organizations.

We think we can be resistant to the temptation of assuming that correlation is actually causation, but it can go against our instinct at times to insist on the distinction. A surprising example of this fell in my lap a few years ago, regarding the health benefits of breast feeding compared to formula feeding young infants. One of the many health claims about breastfeeding is that it reduces the death rate for infants under 1 years old. Most of us who know how vulnerable children can be at this age also know the main causes of death for healthy infants: Sudden Infant Death Syndrome (SIDS) and infections are the first that come to mind -- and indeed these are the two leading causes, followed by injuries. *The New York Times* wrote a controversial news article about the dangers of not breastfeeding, and in it quoted a scientist comparing not breastfeeding to smoking while pregnant [June 13, 2006]. This article was based mostly on the American Association of Pediatrics (AAP)'s claims of benefit.

However, the AAP cited just one study to justify its claim. The statistically significant association between no breastfeeding and infant death rates were immediately convincing of cause for most people. However,,a careful look at the study found that the only cause of death that showed a statistically significant benefit from nursing was *death due to injury* [Chen et al., 2004]. In other words, there was not a statistically significant association between nursing and SIDS, nor between nursing and death due to infection. But there was for injuries. One could still propose that not breastfeeding causes injuries: perhaps the act of holding the baby close to nurse makes the baby less likely to suffer a fatal injury. On the other hand, one is very suspicious that there is a confounder that hadn't been controlled for.

This year the AAP has posted a second study to justify the claim that nursing reduces SIDS, finding a statistically significant association between nursing and less SIDS [Vennemann et al, 2009], though the question of possible missing adjustment for confounders in this analysis still persists given the results on injuries in the previous analysis.

4. Examples in Real News Reporting: Polls

Polls can be horribly misleading if their design and execution are not carefully thought through. A good example of this occurred when The Physicians Foundation conducted a poll, published in November, 2008, and found that 49 percent of primary care doctors planned to cut back or close their practices. As they put it, “The resulting findings show the possibility of significantly decreased access for Americans in the years ahead, as many doctors are forced to reduce the number of patients they see or quit the practice of medicine outright.” A statistician on the poll reported that the margin of error on the poll was one percent, presumably based on how many people responded to the poll.

USA Today headlined their article reporting on the poll, “Primary care doctors in short supply” [November 17, 2008]. The newspaper was distributed around the world. And while the thesis may or may not be true, *the response rate of the poll was four percent*. The possibility of response bias is striking; perhaps those doctors responding to the poll were more likely to be disgruntled doctors or close to retirement. Such a low response rate at least raises a flag: how trustworthy or meaningful are the results?

USA Today did not report on the response rate, as this fact perhaps had little meaning for the author of the article. Certainly, the public was presented with the finding as if it were

“scientific” when to a skeptical eye it provides little to no evidence of a future shortage in doctors.

Perhaps, had the journalist understood the meaning of *margin of error*, the poll would have been less newsworthy. Margin of error only measures how well a sample of poll-respondents represents the whole population (in this case, primary care doctors), *provided that the sample is randomly chosen*. It cannot speak to the bias that could have occurred if the respondents were not random, so in this case, the margin of error has limited applicability. This possible bias should have been disclosed by the group sponsoring the poll, but it may not have been in their interest. Instead, it was up to the media to note that the poll has diminished value because of its poor response rate.

At times, media outlets like to take on their own scientific investigation, and it can have humorous consequences. In the wake of political leveraging before the 2004 presidential election, *Primetime Live* in conjunction with ABC News Polling Unit presented a poll that it had conducted, which found that “More Republicans were satisfied with their sex-lives than Democrats” [Oct. 18, 2004]. Indeed, the poll found a series of amusing details about these folk: 56 percent of Republicans reported being very satisfied with their sex-lives, compared to only 47 percent of Democrats; 72 percent of Republicans wore something to enhance their sex-lives and only 62 percent of Democrats did so; and only 28 percent of Republicans faked an orgasm compared to 33 percent of Democrats.

But the poll did not adjust for a confounding factor: Democrats are more likely to be female than Republicans.

Perhaps journalists reporting on this poll found the results so appealing or humorous that they did not want to investigate whether the poll was really pointing to a difference in politics or simply reflecting gender differences. That said, if news is supposed to be, well, news, then journalists ought to understand what a confounding factor is before they poorly interpret their own polls, leading the public (perhaps with a chuckle) to make conclusions about political parties that may be baseless.

5. The Magnitude of the Problem

One recurring theme is that problems are “declared” without magnitude. It seems that at least once a year there is a report about how kids are doing drugs, drinking and having sex. But how does it compare to previous years? How does it compare to a generation ago? Consistently, we see numbers cited without context: from estimates of environmental damage to discussions of risk.

At times many details are given about the purported problem, with no context for their meaning to non-experts. Recently the *Los Angeles Times* in reference to the recent uncontrolled fire in the Los Angeles area reports “Fighting the Station fire has cost at least \$43.5 million, and federal fire officials say the 154,000-acre blaze in the Angeles National Forest is likely to be one of the most expensive fires in the country this year.” [Sept. 6, 2009] But trying to put those numbers in context is difficult. How many other fires occurred this year? How does it compare to the cost of Katrina or other natural disasters? (the estimates for Katrina are on the order of 200 to 500 times that amount) [*SF Gate*, Sept. 27, 2005]. What proportion of the firefighting budget is that? And how big is 154,000 acres anyway?

Similarly, last year *The New York Times* reported on *dysmenorrhea*, a menstrual disorder characterized by painful cramps in the lower abdomen. Like many "personal" stories about widespread medical problems, it featured a young woman who had the disorder and suffered through it for many years before finally seeking and finding help. It advocates increased awareness of the problem, and points to other medical conditions that painful menstrual periods might indicate, such as endometriosis. But then a side-bar explains that dysmenorrhea is simply "menstrual cramps" in contrast to the article, which calls it "painful cramps." The article notes that this disorder "affects 20 to 90 percent of adolescent girls in some way and severely impacts another 14 to 42 percent." [Nov. 20, 2008]. Leave aside the fact that 90 plus another 14 percent is already over the full population. A range from 20 to 90 percent suggests that dysmenorrhea is extremely ill-defined -- and is so common that it hardly merits the word "disorder."

6. Lacking Evidence but Adding Some Scare

Another common media mistake is the presentation of what looks like "science" without any science behind it, including an acceptance of what might be termed "belief" without any evidence. Discussion of addiction and rehabilitation seem particularly prey to this kind of reporting. A recent article in the *Seattle Post Intelligencer* on what has been termed "internet addiction" referred to the first residential rehab center for the addiction in the U.S. [Sept. 9, 2009]. The article was framed in terms of one boy's story of online gaming and failing out of college. In fact, internet addiction is not recognized as an official term by the American Psychiatric Association. The article contained no discussion of why there is debate among psychologists as to whether self-destructive, excessive computer use should be deemed an "addiction." It tacitly implies that the addiction is a real phenomenon, and that treating it is as scientific as, for example, treating a heroin addiction, which has been studied extensively. Perhaps worse for those who suffer from compulsive computer use, the article did not touch on whether the outpatient rehab centers run by the same people who started the residential one have been effective in helping people reclaim their lives.

At least this article didn't imply that *your* child is going to be an internet addict. In contrast, many stories imply that a problem is so widespread it will likely happen to you or someone near you. They do this without citing any evidence or indication of the prevalence.

"Kids don't have to go very far to find some of the hot drugs of choice — many can find them in their own homes," begins a recent article in the *Seguin Gazette-Enterprise* titled "The Medicine Cabinet Problem" [Sept. 6, 2009]. It would be hard to read this without wondering whether your own kids are exploring the possible pharmaceutical cocktails made from those old muscle relaxers in your medicine cabinet. But while the article spends 1,300 words discussing the legal consequences for kids who sell these drugs and programs designed to help kids realize what they are doing, it never says whether many or few children are (caught) doing this. This illustrates a larger issue: absolute risk -- the risk that something will actually happen to us -- is obscured or ignored in favor of sensational discussions of a "generic" (and seemingly universal) risk.

7. Conclusion

In an era in which Wikipedia and WebMD are considered by experts more reliable than journalism for certain kinds of information, journalists and media sources need to evolve to maintain their relevance. At the same time, journalists are under newer and greater pressures than previously due to budget cuts and shrinking of the news industry. Statisticians can play an important role in this: work with journalists to represent scientific findings accurately and wholly, and encourage them to promote scientific thinking in the mainstream. Statistical literacy is an essential part of life, not just for our students, but also for our media-consuming public.

8. Acknowledgements

Thanks to the W. M. Keck Statistical Literacy Project for inviting this paper for presentation at the Statistical Literacy 2009 session held as part of the 2009 Joint Statistical Meeting of the American Statistical Association.

9. References

- ABC Primetime, *Primetime Live Poll: More Republicans Satisfied With Sex Lives Than Democrats*, October 18, 2004.
- Nicholas Bakalar, *TV Watching in Youth Tied to Depression Later*, in *The New York Times*, February 10, 2009.
- Julie Cart, *Cost of Fighting Station Fire Tops 43 Million*, in the *Los Angeles Times*, September 6, 2009.
- Aimin Chen and Walter J. Rogan, *Breastfeeding and the Risk of Postneonatal Death in the United States*, in *Pediatrics*, Vol 113, No. 5, May 2004., pp. e435-e439.
- Nicholas K. Geranios, *Internet addiction center opens in Fall City*, in *The Seattle Post Intelligencer*, September 7, 2009.
- Ron Maloney, *The Medicine Cabinet Problem*, in *Seguin Gazette-Enterprise*, September 6, 2009.
- Kathleen Pender, *The True Cost of Katrina*, in the *SF Gate*, September 27, 2005.
- Roni Ruben, *Breast-Feed or Else*, in *The New York Times*, June 13, 2006.
- Rita Rubin, *Primary Care Doctors in Short Supply*, in *USA Today*. November 17, 2008.
- Carolyn Sayre, *Taming Menstrual Cramps in Adolescents*, in *The New York Times*, November 20, 2008.
- M.M. Vennemann, T. Bajanowski, B. Brinkmann, G. Jorch, K. Yücesan, C. Sauerland, E.A. Mitchell and the GeSID Study Group, *Does Breastfeeding Reduce the Risk of Sudden Infant Death Syndrome?* in *Pediatrics*, Vol 123, pp. e406-e410.