



Lies, Damned Lies, and Statistics

By Ira Salkin, Ethics Editor

Editor's Note: This article originally appeared in the April 2012 issue of *Editorial Office News*.

Mark Twain popularized a phrase that he probably mistakenly attributed to British Prime Minister Benjamin Disraeli: “There are three kinds of lies: lies, damned lies, and statistics.” This was brought to mind by a recent article that describes the numerous ways that authors may manipulate their data sets so that statistical analyses will support the significance of the results obtained in their studies. In an article entitled “False-Positive Psychology” published in *Psychological Sciences*, Joseph Simmons and colleagues reported that flexibility in data collection methods allows authors to present virtually anything as being statistically significant.¹ The latter can be roughly defined as the opportunity that results can occur by chance rather than by the factors investigated by the researchers as being not greater than five percent.

Simmons and associates point out that it is the researchers (authors) who decide on the size of the study sample, when to stop collecting data, and which outliers in their findings can be excluded from analysis. Often, these researchers define their hypotheses during the course of the investigations rather than at the initiation of the studies. Put another way, the authors determine

their objective(s) on the basis of the results they discover during their ongoing research. By manipulating various aspects of their methods, authors can skew the data such that they can claim statistical significance even if the results are not scientifically supported. One reason for this might be that most editors prefer publishing positive rather than negative findings.

To demonstrate the way investigators can manipulate data, Simmons et al established two test populations of approximately equal size. Each group listened to a different popular song and afterward, they were asked to complete a survey containing a number of variables, including their dates of birth. Surprisingly, the members of the group who listened to song A were found to be over a year younger than those who listened to song B. It should be remembered that no attempt was made to establish membership in either group on the basis of the age of the participants.

How could this be described as a statistically significant result? Well, while Simmons et al reviewed numerous variables obtained in their surveys, they reported on the one that would apparently withstand statistical analysis. When one combines this with the fact that researchers are free to halt the data collection when they choose, one enhances the appearance of the significance of the findings. Simmons reviewed the data in blocks of ten observations or responses to the surveys and stopped the analysis when convenient for his co-authors and himself. Purely by chance, this was the point at which the age gap between the two groups was the highest.

1 Simmons JP, Nelson LD, and Simonsohn U. False-positive psychology. Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* 2011;22(11):1359–66.

Editors all too frequently want to have their journals involved in the publication of “crest of the wave” investigations or studies with unanticipated findings, especially if these are supported by statistical analyses. The inclusion of the latter generally ensures that the results of the studies will be quickly disseminated by the popular media. How many times have you read, heard, or seen the results of clinical investigations, new treatment methods, or chronic diseases being reported as statistically linked to unusual variables, as for example, the age of the participants. Whenever I see that a specific percentage of seniors or Caucasians or Italian-Americans are more prone to acquire a specific disease or can be helped by a specific drug, I remember the fact that 98% of the members of these same groups brush their teeth. Therefore, statistical analysis clearly demonstrates that teeth brushing must be a predisposing factor in acquiring the disease or in the successful treatment with the new drug. Remember that authors are responsible for the selection of the variables to be subjected to statistical analysis, but wouldn't the misuse of statistics be uncovered by other scientists working in the same area of study? The reality is that many editors are reluctant to publish papers that replicate investigations that have previously appeared in the literature. However, there is one caveat to this broad editorial policy generalization—if the confirmatory studies generate as much or more attention grabbing results as the original investigations, editors would very likely publish the report.

Dr. Simmons makes the following suggestions to limit the misuse of statistical analyses by authors who might be submitting their work to your journals:

- Authors must establish specific termination dates for their data collection and describe

them in their work. For example, from the world with which I am most familiar, “Bacteria were grown for 72 hours on the new medium, after which the colonies were observed for morphologic variations.” Seeing this, readers know that all data were collected at a specific time;

- Authors must describe all variables collected during the course of their investigations. For instance, authors should indicate that only the length and width of bacterial cells were measured and were subject to further analysis;
- Authors must describe all experimental conditions, even if discontinued during their investigations. In addition, they should report the justification for eliminating them from the studies. Furthermore, they should discuss the statistical consequences if the conditions had not been terminated. In this way, authors cannot select only those results obtained with techniques and procedures that support their hypotheses.

For a complete discussion of Simmons' recommendations for authors, as well as for reviewers, I refer the readers to the *Psychological Science* website where the entire article can be downloaded at no cost.

The majority of instances of possible statistical manipulation are benign in origin in that the authors did not intend to report fraudulent data. However, ignorance of the appropriate methods seriously compromises the integrity of the findings. Therefore, authors, editors, and editors-in-chief should be educated and vigilant with respect to this critical issue to ensure that papers approved for publication meet valid statistical standards.