Letter to the Editor

The Best Graph May Be No Graph

Dear Editor,

In *CHANCE* 21(2), Howard Wainer writes about "Improving Graphic Displays by Controlling Creativity." He makes good suggestions. In one example (Figure 4), he offers 10 improvements (Figure 5) on a report of "five-year survival rates from various kinds of cancer, showing the improvements over the past two decades" (from the National Cancer Institute). Indeed, the latter figure is neater. But, he missed the most important improvement: not showing the figure in the first place! It's terribly misleading and doesn't necessarily reflect any real improvement "over the past two decades."

The three cancers with survival improvement over time (i.e., breast, prostate, colorectal) are those with intensified screening programs over these two decades. Much, if not all, of the higher survival rates is due to what are called the lead time and length biases of screening. These biases are elementary and fundamental in cancer epidemiology. Lead-time bias is the easier of the two to understand. Someone whose cancer is detected n years early in a screening program lives up to n years longer after her tumor is discovered. The pure bias of n years adds to the cancer survival time of everyone whose tumors were detected by screening. Because of the heterogeneity of cancer, the value of n is highly variable and unknown for any particular tumor. The average of n is also unknown, but it is substantial; it is commonly estimated to be 3–5 years in breast cancer.

The "length" in length bias refers to the tumor's pre-symptomatic period, when it is detectable by screening, called the sojourn time. Aggressive tumors have shorter sojourn times because they grow faster. Indolent tumors have longer sojourn times. Screening finds tumors in proportion to the lengths of their sojourn times. Screening preferentially selects tumors with longer sojourn times and, therefore, tumors detected through screening are slower growing and less lethal. An extreme form of length bias is over diagnosis, in which some cancers are found by screening that would not have caused symptoms or death.

There are many analogues that may help one's intuition regarding length bias, and these should be familiar to statisticians. When you look into the sky and see a shooting star, it's more likely to be one with a longer arc, simply because it's the one you saw. Or, when you select a potato chip from a newly opened bag, it's more likely to be a bigger one, simply because bigger ones are more likely to be selected. Waiting time paradoxes are standard examples. Suppose the interarrival times of buses at a certain bus stop are independently exponentially distributed, all with mean m. You arrive at the stop at an arbitrary time and catch a bus. What is the mean time between the arrival of the bus you caught and that of the previous bus? The answer is $2m$.

I don't mean to suggest that we have not made important strides in treating cancer over the last two decades. We have. But, although Figures 4 and 5 are literally correct, they reflect mostly artifact and greatly exaggerate these strides. Similar figures have been misinterpreted by policymakers and the press and have led to inappropriate recommendations regarding screening, with potentially deleterious effects. The only good use of these figures is as an example for teaching, to demonstrate how easy it is to lie with statistics.

Donald Berry
Head, Division of Quantitative Sciences &
Chair, Department of Biostatistics &
Frank T. McGraw Memorial Chair of Cancer Research,
The University of Texas M.D. Anderson Cancer Center

Howard Wainer responds:

I am delighted that professor Berry raised this issue. While I was preparing this column, I debated with myself (and my colleague, Brian Clauser) this very point and decided not to include it, for it seemed an aside from my main point (fixing graphs to communicate better) and confused the goals of description with those of causation. As a descriptive graph, the figures are correct—survival times *are* increasing. But the causal inference, why they are increasing, is what professor Berry addresses. The issue is how much of the improvement is due to earlier detection and how much is due to improved treatment. This seems to me to be hard to partition. Perhaps by adjusting survival rates by, say, the maturity of the tumor at the time of discovery might provide some help. I would be interested in other schemes that could help us measure the causal effect of the changes in treatment.

Correction

According to Steve Stigler of The University of Chicago, the picture represented on Page 29 of *CHANCE*, volume 21, number 2, is an 1842 posthumous painting of Laplace, not of Chevalier de Mere. We have not located a confirmed image of Chevalier de Mere.