Comment

R. A. Bailey

I congratulate Dr. Chatfield on having the sense and persistence to publish this paper. Its overall message is one of the most important that we should convey to new statisticians. Yet the statistical profession customarily plays down these pitfalls. Why?

Of course, any real experiment has too many idiosyncrasies and pitfalls to be reported in full when it is used only as an example. Nonetheless, I have tried to include instructive pitfalls in my published work. As Chatfield mentions in Section 4, such practicalities and pitfalls are always removed from my papers on the insistence of the editor or the referee, on the grounds that they are uninteresting, irrelevant or uninformative, or that they show what an incompetent statistician I am. Perhaps I am, but if we are not allowed to share experiences of our mistakes how can any of us improve?

Although IDA is important, I would not go so far as to let my final analysis be completely dictated by the IDA. In a well-designed study, data are gathered on the basis of certain questions to be answered and certain models to be assumed. The proposed method of analysis should therefore form part of the protocol, even if it has to be modified later. Indeed, without a proposed method of analysis, how can the statistician advise on data collection?

I agree with the six guidelines in Section 5 and should like to comment on each in turn.

(a) I tell my students that 90% of a statistician's consulting work is asking questions, including questions whose answers seem obvious. The students think that I am exaggerating—until they accompany me to a consultation. I do not think that scientists actively try to mislead us; rather, they try to spare us details which they think will confuse us rather than help. Why do they have this attitude? Why are statistics seen as divorced from context?

(b) In experimental science, the importance of good design of investigations, of adequate procedures for recording data and of verification procedures that allow for human fallibility is generally known, even if all too often ignored. How can the statistician ensure that good data are collected—where I mean "good" rather than "bad," and "collected" rather than "accumulated"—in other situations, especially if her involvement is as a citizen rather than a professional? I am concerned about administrative decisions, often with large financial consequences, that are taken after examination of data that can best be described as "windfall" rather than "collected."

(c) Checking data quality was one of the first things that Dr. Preece taught me when I began work at Rothamsted. Soon afterwards I had to analyse data from a calf-feeding trial. Checking through variables such as birth weight, birth date, weight at eight weeks and amount eaten in certain weeks, I found several inconsistencies in the data and so telephoned the scientist in charge of the trial for elucidation. He told me airily that the data sheets had been filled in, after the end of the trial, by a student who was working for him for a few weeks; he did not seem to think that errors and inconsistencies in his data were any cause for concern. How widespread is his attitude?

(d) When I visited D. Saville and D. Baird at the Biometrics Unit of MAF New Zealand in Canterbury, they took me on some consulting visits and showed me how they refused to analyse any data that the scientist has not already examined graphically. This seems to me an admirable use of IDA, but there are still pitfalls, related to points (a)–(c).

For example, some plant pathologists once showed me six points on an x-y graph and asked me if the points could be said to lie on a straight line. The data came from an experiment on barley. The x values were percentage of tillers infected by leaf-stripe disease; the y values were mean yield in tonnes/ha. The points did indeed appear to lie roughly on a line, but I enquired whether there had been no replication. The reply was that there had indeed been six replications of each dose of infection and that each plotted point represented the mean x-value and the mean y-value for that dose. The plot of all 36 points revealed a very convincing straight line, but with one point lying a long way above the line. At last I asked to see the field note-books. Then I discovered that each x-value was the average of the estimated percentage infection in 10 sampled subareas of the plot. The average was calculated by the field worker, either by

R. A. Bailey is Professor of Mathematical Sciences at Goldsmiths' College in the University of London, New Cross, London SE14 6NW, United Kingdom.