Critical reading and critique of medical articles

To the Editor

Al-Ateeg has given guidance as to how to critically read medical research articles.[1] I will argue below that, more ambitiously, it may be valuable to criticize and critique, especially in respect of the statistical analysis of the results.

Why is the reader reading? To Al-Ateeg, the reader's purpose is to acquire and maintain knowledge. I suggest a reader may have 2 other motivations, which are: 1. Pleasure and relaxation. 2. A hope of being able to actively contribute expertise. I believe that, in the case of papers that give data in sufficient detail, it is common to be able to discover features of the data that were overlooked by the original authors. A reader who is reading critically will often get a degree of pleasure from discovering such features and weighing up what they mean, and will be able to contribute by writing a letter of comment to the relevant journal. Many medical journals have a strong tradition of publication of comments. A distinguished medical statistician and the Editor of the Lancet have both emphasized the importance of this in the scientific communication process.[2,3]

Admittedly, reanalysis of data requires a certain level of statistical knowledge. But I am not suggesting that the reader should check features of the analysis that require advanced statistical concepts. Rather, there are many papers where it is feasible to ask, and then to answer, quite a simple question that was not considered by the original researchers. I have in mind using techniques learned in only a basic statistics course. (As a bonus, the reader's statistical knowledge may deepen through active involvement with a dataset.) Al-Ateeg[1] suggests a reader may work with a statistician in evaluating the data analysis in an article. That collaboration may go further and become a joint exploration of the data.

It is quite common for authors to be too close to their data. They may fail to ask "Is my summary statistic appropriate for its purpose?" or even "Do the numbers capture the concept of interest?" As an example of the first of these issues, correlation coefficients are sometimes calculated when the error of prediction is of greater relevance. As an example of the second, calculations may be carried out on raw numbers when it would be better to first modify the numbers so as to reflect the degree of pathology or danger to life. A reader who comes fresh to the data may be more alert to such issues than the authors were. Hutchinson et al[4] argued that data should be constructed by the active exercise of choice, and numbers should not passively be accepted without thought. Not only the numbers to be analyzed, but also the summary statistic, the null hypothesis, and the alternative hypothesis are choices - and it may be appropriate to challenge the choices made by the authors.[5]

Sometimes, when considering alternative methods of data analysis, it seems that the numbers are unsuited to the most common calculations (such as, of the mean and standard deviation). For example, they may represent ordered categories of outcome or severity, not exact measurements. Questions then arise of whether a nonparametric test should be used rather than a parametric test, or whether the variable should be simplified into a dichotomy. And at this point, a statistician may be consulted. That is appropriate, but often there are more important issues, concerning how the relevant variable should be represented in numbers, and what assumptions are appropriate. These really fall within the clinician's or scientist's field of expertise, not the statistician's. Even in such a case, the statistician may play a useful role, by discussing the issues with the clinician or scientist and helping to make more precise what may begin as quite vague questioning of the authors' treatment of their data.

A point that is often given insufficient consideration is the elementary processing of numbers. Very often, there is a subtraction (such as, change = after minus before) or an addition (such as, when calculating a mean). But a subtraction implies we believe that 8 - 6 is the same as 4 - 2: is a change from 6 to 8 really equivalent to a change from 2 to 4? And an addition implies we believe that 2 + 5 is the same as 4 + 3: is the pair of observations 2 and 5 really equivalent to the pair 4 and 3? The appropriateness of asking these questions is likely to be readily accepted in the context of a crude scale of ordered categories, as (for example) when severity of injury is recorded as 3 = fatal, 2 = serious, 1 = slight, or 0 = none: one death plus one no injury (3 + 0) does not equal one serious plus one slight (2 + 1) injury. But the point is also worth considering with measurements (such as, mm Hg or kg).[4] For example, vehicle speed is a measurement, but the untransformed numbers do not adequately convey the meaning for the safety of a city: the relationship between speed and risk of a crash is strikingly nonlinear, a slight reduction giving a more than proportionate reduction in risk.[6]

If you think that you would enjoy developing your knowledge and making a contribution to the proper interpretation of research results, you need to identify articles that publish the data in detail - for example, as a case-by-case listing, with several variables given for each case. (Datasets of moderate

size, perhaps 20 or 80 cases, are likely to be most suitable: too few, and reliable trends will be unlikely; too many, and the analysis will be too time-consuming.) Some journals have many papers that give data in detail, but others are reluctant to give adequate space for this purpose, though the importance of this has been emphasized.[2] You need a statistical software package to speed the calculations and the plotting of graphs, and it helps if there is a statistical consultant with whom to collaborate. You need a freshness of vision that will permit you to see differently from the authors, along with a sense of balance that will prevent you from giving too much attention to unimportant errors.

Perhaps it is a little fanciful, but I have suggested that spending time on research (and this can take the form of statistical critique of published work) is good for the morale of medical staff.[5] Some hobbies and pastimes are scholarly, and an interest in statistics can be applied there, too. Examples would be easy to find in such fields as mineralogy, nature study, or numismatics.

T. P. Hutchinson
Centre for Automotive Safety Research
University of Adelaide
South Australia 5005
Australia

Reply from the Author

No reply received from the author.

References

1. Al-Ateeg FA. Reading medical articles critically. What they do not teach you in medical school. ***Saudi Med J*** 2004; 25: 409-423.
2. Altman DG. Poor-quality medical research. What can journals do? ***JAMA*** 2002; 287: 2765-2767.
3. Horton R. Postpublication criticism and the shaping of clinical knowledge. ***JAMA*** 2002; 287: 2843-2847.
4. Hutchinson TP, Cairns D, Chekaluk E. The construction of data to reflect the research objective, and how randomisation tests make such data usable. ***Statistical Papers*** 2002; 43: 349-359.
5. Hutchinson TP. Statistics — for fun and therapy. ***Br J Gen Pract*** 2004; 54: 228-229.
6. Kloeden CN, McLean AJ, Moore VM, Ponte G. Travelling speed and the risk of crash involvement. Volume 1: Findings. Canberra: Federal Office of Road Safety; 1997. Report CR 172.