sets of data can be described exactly by the means and SEs in fig 2. This is because fig 2 tells us nothing about differences between machines for each subject. Error bars are thus useless in the case of paired observations.

Now suppose that we wish to compare the diastolic blood pressures of two distinct groups of people, say doctors (group A) and bus-drivers (group B). Figures 4(a) and 4(b) show two possible outcomes. In which case, if either, are the two groups significantly different? It is not easy to tell from the raw data shown that the groups are significantly different in fig 4(a) (p <0.05) but not in fig 4(b) (p >0.1). What would an "error-bar" plot show? Well, again both examples would yield fig 2, showing that the visual impression of non-overlapping bars does not by itself give any information about statistical significance. If the error bars do overlap, however, then the difference between the means is not statistically significant.[11]

For error bars to be useful they ought to convey useful information about either the precision of individual means or the differences between means. In their usual form they do neither, although my impression is that many people believe that they do both. The use of confidence intervals (mean± 2 SE) instead of error bars does at least give useful information about individual means. Although it is sometimes possible to make the visual presentation give an indication of statistical significance, it is probably better to give confidence intervals and, if desired, report on the significance separately.
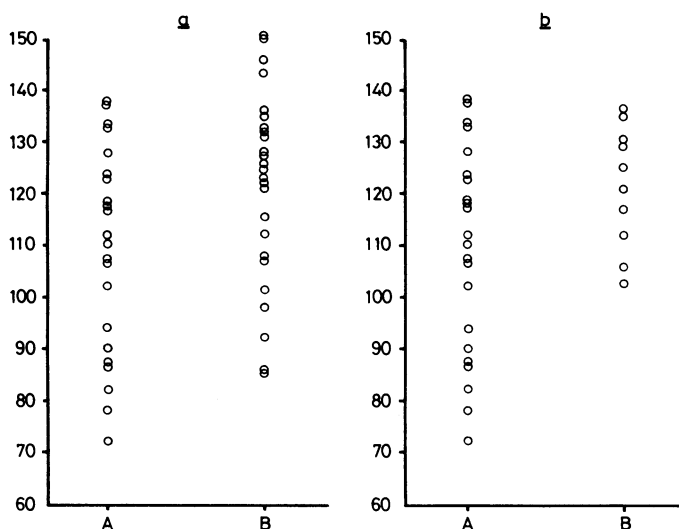

FIG 4 (a) and (b)—Comparisons of diastolic blood pressure in two different groups of subjects.

Numerical precision

One other aspect of presentation that deserves some comment is numerical precision. It is rarely necessary to quote results—means, standard deviations, and so on—to more than three significant figures (that is, excluding leading or trailing zeros). For tabular presentation it may be a positive advantage to reduce the precision of each entry to make any patterns or trends more obvious.[12]

Spurious precision should also be avoided. Examples are the quoting of t or χ^2 values to four decimal places, and a regression slope with seven significant figures (12·97642). My favourite is the summary[13] of a test of significance as p $<10^{-54}$, although I must concede that there is only one significant figure here!

Some suggestions

More thought should be given to numerical and visual presentation, rather than automatically following precedent.

Some ways of supplying more information without using more space are:

(1) In a plot information about the spread of data (by ± 2 SD ranges or centiles) can be given as well as means and confidence intervals.

(2) A figure and a table may be combined by using the X axis labels as table column headings. For example, in fig 2 I could have given the mean, SD, range, and sample size for the two groups under the figure using little extra space.

(3) When scatter plots have the same variable on each axis as in fig 3(a) and 3(b), a small histogram of the within-person differences can be added in an otherwise empty corner.

Summary

Whatever results are presented it is vital that the methods are identified. In one survey of over 1000 papers[14] as many as 20% of the procedures were unidentified, and in another it was not clear whether the SD or SE was given in 11% of 608 papers.[6] It is impossible to appraise a paper in the presence of such ambiguities.

Visual display is a particularly effective way of presenting results. Given alternatives, however, many people might opt for the method of display that fits in better with their beliefs. If decisions are taken as a result of such presentations then there is scope for manipulating events by choice of presentation. This practice is well recognised in the way statistics are sometimes presented in the mass media and advertisements; we should not rule out this phenomenon in the medical world.

This is the sixth in a series of eight articles. No reprints will be available from the authors.

References

[1] Huff D. *How to lie with statistics.* Harmondsworth: Penguin, 1973.
[2] Department of Health and Social Security. *Prevention and health: everybody's business.* London: HMSO, 1976.
[3] Radical Statistics Health Group. *Whose priorities?* London: Radical Statistics, 1976.
[4] Armitage P. *Statistics in medical research.* Oxford: Blackwell, 1971:150-6.
[5] Hill GL, Bradley JA, Collins JP, McCarthy I, Oxby CB, Burkinshaw L. Fat-free body mass from skinfold thickness: a close relationship with total body nitrogen. *Br J Nutr* 1978;39:403-5.
[6] Bunce H, Hokanson JA, Weiss GB. Avoiding ambiguity when reporting variability in biomedical data. *Am J Med* 1980;69:8-9.
[7] Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. *Circulation* 1980;61:1-7.
[8] Gardner MJ. Understanding and presenting variation. *Lancet* 1975;i:230-1.
[9] Mainland D. SI units and acidity. *Br Med J* 1977;ii:1219-20.
[10] Feinstein AR. Clinical biostatistics. XXXVII Demeaned errors, confidence games, nonplussed minuses, inefficient coefficients, and other statistical disruptions of scientific communication. *Clin Pharmacol Ther* 1976;20:617-31.
[11] Browne RH. On visual assessment of the significance of a mean difference. *Biometrics* 1979;35:657-65.
[12] Ehrenberg ASC. Rudiments of numeracy. *Journal of the Royal Statistical Society* Series A 1977;140:277-97.
[13] Vaughan Williams EM, Tasgal J, Raine AEG. Morphometric changes in rabbit ventricular myocardium produced by long-term beta-adrenoceptor blockade. *Lancet* 1977;ii:850-2.
[14] Feinstein AR. Clinical biostatistics. XXV A survey of the statistical procedures in general medical journals. *Clin Pharmacol Ther* 1974;15:97-107.

Correction

Evaluation of a patient education manual

The authors of this paper (4 October, p 924) wish to apologise for inadvertently failing to acknowledge the important contribution of Dr Mick Murray in constructing the questionnaires used in the study and his advice on the study design.