Numerical results: some hints on presentation

T D V SWINSCOW

Two fairly distinct kinds of articles with numerical data appear in medical journals. The aim of the first is primarily educational. The author is not putting forward any original, novel, or unusual material. He is trying to present his ideas in the most memorable way possible. In this type of article diagrams and graphs often serve his purpose well by embodying his message in an easily remembered picture. The hints on presentation to be given here are concerned with the second type of article. Its aim is to record a discovery in such a way that the facts and the inferences drawn from them can be examined with precision from every viewpoint and withstand all reasonable criticism of their validity. Too often scientific results in the second type of article are presented in a manner more appropriate to the first type of article. Here the following passage from Bradford Hill[1] is worth bearing in mind.

"Graphs should always be regarded as subsidiary aids to the intelligence and *not* as the evidence of associations or trends. That evidence must be largely drawn from the statistical tables themselves. It follows that graphs are an unsatisfactory *substitute* for statistical tables."

One of the commonest questions asked is, How much detail should I give in my numerical results? An investigator has perhaps measured the blood pressures of 100 patients, given them a hypotensive drug for three months, and measured their blood pressures again to see what benefit has come from the treatment. Should he give all the recordings made on both occasions? Or the mean (average) blood pressures alone? Or the mean plus some measure of dispersal such as the standard deviation? No fixed rule can be offered, but this question needs to be considered afresh in every case with the following idea in mind: the object is to present the observations in enough detail to allow the reader to examine critically the inferences drawn from them yet sufficiently summarily to bring out their meaning. The reader does not want either a profusion of numbers looking like a computer print-out or a stark total that has swallowed up all the details. Such data are often adequately summarised in a frequency distribution with mean and standard deviation. Too often they appear on the page in the form of a histogram which is impossible to read precisely.

Tables

When preparing a table an author should carefully consider its exact purpose, for a table provides an analysis of the data. It always classifies data, often summarises them, and should present them in a logical arrangement that allows a reader to make comparisons or draw inferences from them easily. In a scientific paper the author should always give a table a caption stating its contents, and then explain the contents more fully in the text. Needless to say, figures in tables should be added up correctly and agree with those given in the text. This elementary advice is mentioned only because readers would be amazed if they knew how often conscientious editors find simple errors of this kind in papers when preparing them for publication.

Some arithmetical tact is desirable when results are given in numerical form. For instance, there is no reason why numbers smaller than 100 should not be presented as percentages, but the following figures given in a respected specialist journal show how this can be overdone: "Operations 7, thrombosis 1 (14%), haematoma 6 (86%), dehiscence 0 (0%)." More in the same vein followed. Such percentages exemplify a common fault in medical journals, though it is a failure of tact rather than a lapse into error. A lack of tact is also displayed when the author presents the mean of some observations in the form $2{\cdot}5 \pm 1{\cdot}2$. What does the $\pm$ signify? It is an imprecise notation applied to both the standard deviation and the standard error. A better procedure is to say 2·5, SD 1·2, or 2·5, SE 1·2. But if an author does use $\pm$ he should clarify its meaning. Something worse, a logical error and a common one, annoys readers when an author presents his results in categories that are not mutually exclusive, such as age groups 10-15, 15-20, 20-25, and so on; yet the original records were probably made correctly.

If numerical data are presented to allow a comparison of some sort to be made, a statistical test of the significance of one or more differences between them is nearly always advisable. For instance, a paper showed that 43 patients out of 50 (86%) improved on treatment A while only 20 out of 30 (67%) improved on treatment B. The authors simply left it at that, but a χ^2 test with Yates's correction shows that the difference is not significant at the 5% level. The authors ought to have carried out this or a similar test because their results cannot be reliably interpreted without it. But, if they had, they would then have had the problem of telling their readers what practical inference to draw from it. Some might take the risk of saying that though the difference is not significant at the 5% level it nevertheless indicates a trend that has a bearing on clinical practice. Some would not take that risk because even a 5% level is not at all stringent. Nor is it justifiable simply to add some more cases in the hope of getting a definite answer. Authors who want to publish data that have not reached a conventional and quite low level of statistical significance need to exercise the greatest care in drawing conclusions from them and should usually seek the guidance of an expert statistician.

Statistical help

But even an expert statistician may fail to give the right kind of help if he is consulted too late or advised inadequately.

The old jibe that "statistics can prove anything" owes some of its force to those mistakes. For example, a team of investigators have produced a mass of numerical data and they ask a statistician to make all sorts of comparisons between them. Perhaps he does 20 tests and only one of them is significant at the 5% level. On chance alone this is not unexpected (as he will probably point out), but the investigators use it in their paper, albeit with a slightly disconsolate air, to prove that X is a better tranquilliser than Y, at least for unmarried women with red hair, aged 30-32, whose occupation is personal assistant and whose water supply contains less than 1 ppm of magnesium.

A fallacy of a similar kind can lie in the production of a significant result from an intricate statistical analysis when a simple method applied to a straightforward comparison has failed to give one. It is true that the intricate method may have succeeded because it has made use of more information than the simple one and thereby allowed a more complete comparison. But its success may also be due to the fact that some of the information it uses has little or no importance in clinical practice.

Many observations made in clinical medicine are of necessity

British Medical Journal, London WC1
T D V SWINSCOW, MSC, MB, deputy editor

approximate in their measurement, variable from time to time, and responsive to the investigator's attitude and skill. The application to them of complex statistical methods can easily lead an investigator to draw conclusions that are more precise or refined than the data can properly bear. Two points are worth remembering here: it is far more difficult to make an exact observation than to do a simple statistical test; and statistical significance does not necessarily imply clinical significance.

The difficulty of making an exact observation is matched in clinical practice by the difficulty of making a valid comparison. Statistical methods allow to a certain extent for the averaging out of differences, and in the selection of individuals to make up groups for comparison the use of tables of random numbers is now standard practice. But in reporting the results the investigator should state clearly the composition of his groups so that readers can judge their comparability. Thus it is far too casual merely to say that two groups of patients were comparable in all important respects. Among the attributes in which groups of patients are commonly required to show similarity are age, sex, and socio-economic class or occupation, together with many others in special circumstances, such as family history, personal history, previous exposure to an infection, inoculation state, severity of disease, type of pathological lesion, and prognosis. These attributes or whatever others are considered important to the outcome of the investigation should be clearly set out for the reader's inspection.

Though statistical analysis of clinical data is generally best restricted to simple methods, much ingenuity and subtlety may be needed in planning the investigation that is to provide the data for the analysis. Consequently before presenting the numerical results an investigator should describe the plan of research clearly, for besides setting the results out in numerical form his task is to convey their full and exact meaning. Thus the reader wants not merely numbers but the practical realities that they measure.

Reference

[1] Hill, A B, *Principles of Medical Statistics*, 9th edn. London, Lancet, 1971.

Writing the MD thesis

CLIFFORD HAWKINS

The degree of Doctor of Medicine (MD or DM) in the United Kingdom is a higher doctorate which is coveted by medical graduates, being equal in status to doctorates in other university faculties; this contrasts with practice in countries elsewhere in Europe and in the USA, where it is solely a qualifying degree like the MB, BS. The award of Master of Surgery (MS or in Latin ChM) has the same status and requirement; so much so that this degree was discontinued in the University of Birmingham in 1974, the MD subsequently being awarded to both physicians and surgeons, and this has happened elsewhere. Research interests of both disciplines are similar—a far cry from the days when the candidate for the MS was required to prove his ability at anatomy and the MD degree was awarded by examination, like a bar to the Membership of the Royal College of Physicians examination (MRCP). A doctorate of philosophy (PhD) is also awarded to medical graduates, though it tends to have a lower status, as it is the result of supervised research often obtained at the start of a career. The use of the word philosophy for a science degree is anachronistic: it dates back to the original meaning of philosophy, which covered wisdom and knowledge generally.

The MD thesis is a test more of scientific rather than clinical ability and normally provides some contribution to medical knowledge. The most important quality needed for undertaking it is enthusiasm for original work and for studying a subject in depth. Anyone can present a thesis to their university though obviously time and research facilities are needed, yet some to their great credit obtain it from general practice.[1] It also comes easier to those who have already written articles and received the criticisms of editors.

Anyone intending to do an MD thesis should visit his medical school library and peruse theses already accepted. A room is often devoted to these, all standardised with a similar immaculate binding except for colour. Our university has a different colour for each faculty—for example, the binding is red for medicine and grey for science, and these colours correlate with those on the academic gowns. Some are big and others are slender; one in our library, perhaps a record, weighs 5 kg (10 lb) and consists of three volumes. The smallest is 1 kg (2 lb). Size, however, is no guide to quality; indeed, great length may be due to literary incapacity. A first-class thesis is often small and concise.

Looking at the titles of MD theses will show the wide range of topics, varying from investigation of clinical conditions such as Raynaud's syndrome to experimental work such as the teratogenic action of trypan blue. Other material can be used to support the thesis: audiovisual, tape, gramophone record, a book that the author may have published, or a computer printout.

Choosing a subject

Professional advice must be sought whether a subject for research is viable; else a candidate may easily set out on a task which is too ambitious. A suitable person may be at hand; otherwise an approach should be made to the head of the appropriate department, such as medicine, surgery, obstetrics, psychiatry, oncology, immunology, geriatrics, pharmacology, social medicine, and so on. This is particularly important for anyone working in isolation from a main university department, as discussion must, from the candidate's interest, take place before work is far advanced, and certainly before the thesis is written. If necessary, the dean of the medical school can be asked to suggest someone. Choice of a subject is easiest when the person is one of a research team with an ongoing programme —provided he does the work himself.

The topic must contain sufficient material for a thesis and originality is important, though an excellent piece of work confirming what is known would sometimes be acceptable. There must be a consistent theme. A thesis that is disjointed will be turned down, so that it is no good pinning together work already done on a few vaguely related subjects.

As work progresses it may be given as a communication from time to time at medical meetings. A critical audience may be of great help, especially when they make constructive suggestions. Similarly, any opportunity to publish work should be seized. If accepted, an article appearing in a reputable medical journal will act as an incentive. Reprints can be inserted in a special pocket at the end of the thesis, or copied by photostat and made

Queen Elizabeth Hospital, Birmingham, B15 2TH

CLIFFORD HAWKINS, MD, FRCP, consultant physician and senior clinical lecturer to the University of Birmingham