Medical Journal of the
Islamic Republic of Iran

Volume 8
Number 2
Summer 1373
August 1994

ETHICAL ISSUES IN MEDICAL STATISTICS

SEYYED MOHAMMAD TAGHI AYATOLLAHI, Ph.D., FSS, C. Stat.

From the Department of Biostatistics, Shiraz University of Medical Sciences, Shiraz Islamic Republic of Iran.

ABSTRACT

Medical statistics (biostatistics), as a vital essential part of modern life, does raise some fundamental ethical issues. Surprisingly, this aspect seems to have been totally ignored by books on medical ethics. This paper discusses how the statistical aspects affect the ethics. The relation between biostatistics and medical research is explored. All stages of a medical research exercise are vulnerable to statistical mismanagement which might lead to misuse of patients by exposing them to unjustified risk and inconvenience; the misuse of resources including the researchers' time, which could be better employed on more valuable activities; and the consequences of publishing misleading results, which may include carrying out unnecessary further work. These are specific and highly undesirable outcomes. Failure to guard against these is surely as unethical as using experimental methods that offend against moral principles, such as failing to obtain full informed consent from subjects. Raising statistical standards of medical researches and publications serves as a safeguard to observe the element of ethics. This can be achieved by widespread teaching of medical statistics at all levels, involvement of biostatisticians as active participants of any medical researches and ethical committees. Ethical issues in medical statistics require wider and more open debate. Those involved in medical research need to involve the whole medical profession. Indeed, moral philosophers, theologians, and other professional groups have an important contribution to make.
MJIRI, Vol. 8, No. 2, 121-125, 1994.

INTRODUCTION

Statistics is the science of collecting, summarizing, presenting, and analyzing data. This analysis may lead to conclusions and subsequent decisions. The science of medical statistics (biostatistics) embraces those techniques pertaining to the medical field. Although the methodology of statistics is quite general, specific problems arise in medicine, often because the unit of interest is a living person rather than some abstract phenomenon, object, or financial consideration. The following questions, some of which are faced nearly every day by the practising physician, are to a great extent statistical in nature: Is this new drug or procedure better than that commonly in use? How much better? What, if any, are the risks or side effects associated with its use? In testing a new drug how many patients must be treated, and in what manner, in order to demonstrate its worth? What is the normal variation in some clinical measurements? How reliable and valid is the measurement? What is the magnitude and effect of laboratory and technical error? How does one interpret abnormal values?

In fact, the two simple questions a patient can ask his physician-namely: "What's the diagnosis?" "What are my chances, Doctor?" -are very frequently statistical in nature. In order to seek a sound answer one often relies on the results of statistical analysis. This is why Francis Galton' speaks of statistics as follows:

> "Some people hate the very name of statistics but I find them full of beauty and interests. Whenever they are not brutalized, but delicately handled by the higher methods, and are warily interpreted, their power of dealing with complicated phenomena is extraordinary. They are the only tools by which an opening can be cut through the formidable thicket of difficulties that bars the path of those who pursue the Science of man"

Such problems and questions, which are linked to issues merit special attention. Underscoring the need or greater understanding of statistics in medicine is the fact that much of statistical material in the medical literature is improperly conceived, executed, or interpreted, which is obviously unethical. Surprisingly, this aspect seems to have been totally ignored by books on medical ethics. Thus, the purpose of this paper is to discuss briefly a much neglected aspect of medial ethics-how the statistical aspects affect the ethics. To do so, we follow the structure of a research exercise, outlined in the following figure:

PLANNING
↓
DESIGN
↓
EXECUTION
(data collection)
↓
DATA PROCESSING
↓
DATA ANALYSIS
↓
PRESENTATION
↓
INTERPRETATION
↓
PUBLICATION

Relationships Between Biostatistics and medical ethics

The very first question that one can ask is: What is the relation between medical statistics and medical ethics? It is well appreciated that ethical considerations may affect the design of an experiment. Perhaps the most obvious examples are clinical trials-we cannot, for example, carry out controlled trials of cigarette smoking. Stated simply, it is unethical to carry out bad scientific experiments.[2] Statistical methods are one aspect of this. However praiseworthy a study may be from other points of view, if the statistical aspects are substandard then the research will be unethical. There are two principal reasons for this.

Firstly, the most obvious way in which a study may be deemed unethical, whether on statistical or other grounds, is the misuse of patients (or animals) and other resources. As May[3] has said:

> "... one of the most serious ethical problems in clinical research is that of placing subjects at risk of injury, discomfort, or inconvenience in experiments where there are too few subjects for valid results, too many subjects for the point to be established, or an improperly designed random or double-blind procedure."

Secondly, however, statistics affects the ethics in a much more specific way: it is unethical to publish results that are incorrect or misleading. Errors in the use of statistics may occur at all stages of an investigation, and one error can be sufficient to render the whole exercise useless. A study may have been perfectly conceived and executed, but if it is analyzed incorrectly then the consequences may be as serious as for a study that was fundamentally unsound throughout.

Study Design and Sample Size

In observational studies data from a sample of individuals are used, either implicitly or explicitly, to make inferences about the population of interest, such as men aged 20-65, hypertensive, or pregnant women. The major problem of all observational studies is the selection of subjects for study. This aspect must be given considerable attention at the design stage, because if the sample is not representative of the population then the results will be unreliable and of dubious worth.

Clinical trials of some sort are clearly important for new treatments. As May[3] says:

> "The ethical justification for such experimentation, which is outside the pure physician-patient relationship, is based on a judgement that in certain circumstances it is legitimate to put a subject at risk, with his or her consent, because of the overriding need of society of progress in combating certain diseases".

The debate about the ethics of clinical trials is still very active. Some authors have suggested that it is unethical not to carry out a clinical trial on a new treatment, whereas others believe that such trials are unethical, at least in the way they are usually conducted.[4]

There is no one best design for all clinical trials. The choice for a specific trial must depend on the seriousness of the conditons being treated, the nature of the treatments, the response time, the measures of outcome, and so on. The

main ethical problem is balancing the interests of the individuals in the study with those of the much larger number who may benefit in the useful results, and this may often be achieved best by a randomized study (double-blind if possible). If it is thought likely that highly favourable early results or a high incidence of side effects would argue in favour of premature termination of the study, then these considerations may be built in, using a sequential design.

The ethical difficulties associated with the widespread use of a new treatment without a trial are far greater than those with the trial itself. The importance of good design, however. is reflected in the many examples of conflicting results that may be found in a series of case-control studies of the same topic.[5] As a notable example, after 32 studies over 25 years there is still no consensus on the efficacy of anticoagulants following myocardial infarction.[6]

Whatever type of statistical design is used for a study, the problem of sample size must be faced. This aspect, which causes considerable difficulty for researchers, is perhaps the most common reason for consulting a statistician.

The idea behind using the concept of power to calculate sample size is to maximise, so far as practicable, the chances of finding a real and important effect if it is there. and to enable us to be reasonably sure that an negative finding is strong for believing that there is no important difference.

Before embarking on a study the appropriate sample size should be calculated. If not enough subjects are available then the study should not be carried out or some additional source of subjects should be found.[7] (It should also be borne in mind that expected accession rates tend to be over-optimistic). The calculations affecting sample size and power should be reported when publishing results. A study[8] of 172 randomized controlled trials published in the New England Journal of Medicine and the Lancet from 1973 to 1976 found that none mentioned a prior estimate of the required sample size, and none specified a clinically relevant difference that might allow calculation of the power of their study. Obviously in most of these studies such calculations were not done.

It is surprising and worrying that in such an ethically sensitive area as clinical trials so little attention has been given to an aspect that can have major ethical consequences. If the sample size is too small there is an increased risk of false negative finding. A survey[9] of 71 supposedly negative trials found that two-thirds of them had at least a 10% risk of missing a true improvement of 50%. In only one of the 71 studies was power mentioned as having been considered before carrying out the study. It is surely ethically indefensible to carry out a study, with only a small chance of detecting a treatment effect unless it is a massive one, and with a consequently high probability of failure to detect an important therapeutic effect.

EXECUTION: COLLECTING AND SCREENING DATA

Problems with data collection are often the result of the failure at the design stage to anticipate unusual circumstances. This is one reason why large studies ought to have a pilot phase to try to spot any major deficiencies. It is because we cannot foresee everything that may be relevant that randomization is so important, but it must be strictly adhered to.

The issues of data screening generally receive scant attention. Yet they concern strategic decisions that can have major implications for the ensuing results, as the criticism[10] of Anturane study[11] has shown. They directly affect the validity and thus the ethics of research.

DATA PROCESSING

The wide availability of computers and calculators has made it much easier to carry out statistical analyses. Unfortunately, they have also made it easy to produce results without ever really studying the raw data. Before embarking on analysis there is much that can be learnt from simple inspection of variables both singly and in pairs. Such screening of the data, especially graphically, as well as greatly helping to prepare the data for analysis, can also provide considerable insight into the relationships between variables.

ANALYZING DATA

The incorrect analysis of data is probably the best known misuse of statistical methods, largely due to a series of reviews[12] that have shown how common such errors are in published papers. Nevertheless. these mistakes, which tend to be in the use of the simpler techniques, continue to proliferate. The mishandling of statistical analysis is as bad as the misuse of any laboratory technique. Both can lead to incorrect answers and conclusions and are thus unethical because they render research valueless.

It is of no value collecting good data if the analysis is inadequate or invalid. The results obtained may then be worthless, or at best they will fail to realize the true potential of the data. Either way, the value of the whole experiment is diminished to a point where the ethics of the investigation must be called into question.

PRESENTATION OF RESULTS

A very important aspect of statistical method is the clear numerical and graphical presentation of results. It is

uncommon to find discussion of how best to present the results of statistical analyses. This is surprising, since the interpretation of the results, both by the researcher and by later readers of the paper, may be critically dependent on the methods used to present the results.

Whenever results are presented it is vital that the methods are identified. In one survey of over 1000 papers[13] as many as 205 of the procedures were unidentified, and in another it was not clear whether the standard deviation (SD) or standard error (SE) was given in 11% of 608 papers.[14] 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 of manipulating in the way statistics are sometimes presented in the mass media and advertisements; we should not rule out this phenomenon in the medical world.

INTERPRETING RESULTS

Pearson and Hartley[15] explain the role of statistical method as follows:

> "... it is a function of statistical method to emphasise that precise conclusions cannot be drawn from inadequate data".

Some errors are specific to the interpretation of results. Most emphasis should be given to tests of significance, since these quite clearly cause great difficulty.

The enormous amount of published research makes it inevitable that papers will often be judged, in the first instance at least, by the authors' own conclusions or summary. It is thus vitally important that these contain valid interpretations of the results of the study, since the publication of misleading conclusions may both nullify the research in question and falsely influence medical practice and further research.

PUBLICATION

Once published, a piece of research achieves both respectability and credibility so that it is important for journals to make strenuous efforts to detect substandard research. In recent years there have been several good studies of the quality of statistics in papers in medical journals to support the idea that there is much room for improvement. For example, Schor and Karten[12] reported that of 149 papers reporting analytical studies in several journals, only 28% were judged acceptable, 67% were deemed deficient but could be improved and 5% were totally unsalvageable.

The editor of the journal wrote as follows[16]:

> "The study is an indirect argument for greater knowledge and appreciation of statistics by the medical author, for a reiteration on his part that the biostatistician is not a worrisome censor, but a valuable ally, and that biostatistics, far from being an unrelated mathematical science, is a discipline essential to modern medicine-a pillar in its edifice".

The ethical implications of publishing research containing incorrect or unfounded results or conclusions are little affected by the nature of the errors made, and are indeed much the same as the consequences of publishing spurious results. The cost in time and energy in trying to reproduce such results can be enormous.[17] Alternatively, the results may rest unchallenged for many years. Suppose a randomized controlled trial is carried out in which a conclusion is reached that the new treatment is significantly better than the previous standard treatment. The publication of such a finding may well affect patient care, and it may then be considered to be unethical to carry out further trials as one group would be denied the new treatment that was "known" to be better. Clearly, both of these consequences of publication will hold whether or not the conclusions were justified unless any deficiencies are very obvious or if there is considerable protest. A solitary critical letter, perhaps from a statistician, hidden away on the correspondence page is unlikely to be efficient. Similar consequences apply in the opposite case where a treatment is incorrectly found to be ineffective.[18]

CONCLUSIONS

By emphasizing the ethical implications of carrying out research and publishing papers with incorrect statistics, the time has arrived to raise statistical standards in medical research to observe elements of ethics and benefit mankind. The immediate needs to achieve this goal are:

a) Teaching of biostatistics

The recent widespread move to include biostatistics in the syllabus for medical and paramedical students in Iran is a welcome development. Such teaching is likely to be most beneficial when it gets away from rigid method-oriented approach and concentrates more on general concepts. For medical students it may be more successful when not taught as an isolated subject, but closely related to another course such as epidemiology.[19]

An aid is provided for clinicians and medical investigations who are planning their continuing education in statistical methods, and faculty who design or teach courses in quantitative methods for medical and health

professionals.[20]

Of greater value in this respect would be postgraduate courses in statistics for those who had previously had an introductory course, and aimed particularly at those intending to do research. Such courses should try to give a greater understanding of statistical concepts: to help researchers to understand properly the simpler statistical methods (including when not use them), to appreciate the principles of more advanced methods, and to know when to seek expert help.

b) Involvement of medical statisticians

In general, the larger a project the more likely it is that a biostatistician will be directly concerned. Unfortunately, not all medical researchers have direct access to medical statisticians, but large collaborative studies usually need considerable statistical advice, preferably with a biostatician as an active participant. Even for small studies statistical advice before the research beginning may be valuable, especially in helping to match the design to greater understanding of the research. Yet, despite the common plead for early involvement, most consultancy concerns the analysis of data that have already been collected. A bigger problem though is that many projects are carried out without the benefit of any statistical advice at all. Increased involvement of medical statisticians in medical research would clearly improve the overall standard of statistics, but this requires greater availability of biostatisticians than present.

Successful consultancy relies on the ability of both researcher and statistician to understand each other's language, which is not always easy.

c) **Ethical committees**

Establishment of active ethical committees would provide the opportunity to review many protocols for intended research on human subjects, and have the important sanction of withholding their approval. In view of the ease with which research can be rendered unethically statistical mismanagement it should be an automatic part of the review by ethical committees to look formally at the experimental design, and preferably also at the intended form of analysis. Medical statisticians should be represented on ethical committees actively.

REFERENCE

1. Galton F: Natural Inheritance. London: Macmillan, 1889.
2. Denham MJ, Foster A, Tyrrell DAJ: Work of a district ethical committe. Br Med J ii: 1042-5, 1979.
3. May WW: The compostion and function of ethical committees. J Med Ethics 1: 23-9, 1975.
4. Pocock SJ: Clinical Trials-A Practical Approach. Chichester; John Wiley & Sons, 1989.
5. Horwitz BI, Feinstein AR: Methodological standards and contradictory results in case-control research. Am J Med 66: 556-64.
6. Doll R, Peto R: Randomized controlled trials and retrospective controls. Br Med J 280: 44, 1979.
7. Peto R, Pike MC, Armitage P, et al: Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 34: 585-612, 1976.
8. Ambroz A, Chalmers TC, Smith H, Schroeder B, Freiman JA, Shareck EP: Deficiencies of randomized control trials. Clin Res 26: 280A, 1978.
9. Freiman JA, Chalmers TC, Smith H, Kuebler RR: The importance of beta, the type II error and sample size in the design and interpretation of randomized control trial. N Engl J Med 299: 690-4, 1978.
10. Kolata GB: FDA says no to Anturane. Science 208: 1130-2, 1980.
11. The Anturane Reinfarction Trial Research Group. Sulfinpyrazone in the prevention of sudden death after myocardial infarction. N Engl J Med 302: 250-6, 1980.
12. Schor S, Karten I: Statistical evaluation of medical journal manuscripts. JAMA 195: 1123, 1976.
13. Feinstein AR: Clinical biostatistics. XXV, A survey of the statistical procedures in general medical journal. Clin Pharm Ther 15: 97-107, 1974.
14. Bunce H, Hokanson JA, Weiss GB: Avoiding ambiguity when reporting variability in biomedical data. Am J Med 69: 8-9, 1980.
15. Pearson ES, Hartley HO: Biometrica tables for statisticians. Volume 1, third edition. Cambridge: University Press, 83, 1970.
16. Anonymous: A pillar of medicine. JAMA 195: 1145, 1966.
17. Muller M: Why scientists don't cheat: New Scientist, 74: 522-3, 1977.
18. Altman DG: Misuse of statistics is unethical. Br Med J 281: 1182-4, 1980.
19. Ayatollahi SMT: Medical Statistics. Teachers' Meeting. Bristol, 1993.
20. Emerson JD, Golditz GA: Statistics in practice-use of statistical analysis in the New England Journal of Medicine. N Engl J Med 309: 709-13, 1983.