The Usage of Statistics in Medicine

A S Atiya, MSc, Department of Social & Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur

Summary

Over the years there is an increasing trend in statistics usage including the use of sophisticated techniques in medical journals. While its use should be encouraged, one ought to use it wisely, lest the study findings may be invalid. One of the main reasons for its widespread use is the availability of statistical softwares in the market. There is a tendency for doctors to use it without basic understanding of its use. This article first in the series of continuing medical education in medical statistics attempts to highlight some of these issues and offer some remedial measures that can be taken in the prevention of its misuse.

Key Words: Understanding of statistics, Medical education, Statistical programmes

Introduction

The increasing use of statistics in medical journals is well documented[1]. There are well over 20,000 journals catering to the health care community and majority contain articles that require testing of hypothesis. What is even more alarming is the increasing use of sophisticated statistical analysis. A pertinent question that needs to be raised is what prompted todays' investigators to use such analysis. Is it because the research questions posed demand beyond the use of *t*- and Chi-square tests of significance? Or is it its use will increase the chance of publication? The answer to these questions is not a straight forward one.

Beside the actual need to carry out research, the proliferation of such interest among medical doctors may be attributed to the availability of computer softwares. Over the last decade there are numerous statistical packages, e.g. SPSS, SAS, MINITAB, SYSTAT, and INSTAT just to name a few, that can easily be purchased from the vendors. Some of them are so user friendly that an unwary person may easily be seduced to use it - and why not? Today, even few students would possess a copied version of some elementary statistical package. These are some of the main reasons for its widespread use. These reasons are however general observations.

Even if such answers are readily available, the more important consideration is whether the investigators know the justification of its use. Echoing what has already been said in numerous citation - misuse of statistics is unethical[2]. The pillar of good medical practice is that one must be consciously aware of ethics in whatever one does in relation to his patients. Carrying out a poor piece of research is no different to that of incorrectly instituting a wrong prescription to the patients.

It is with the above notion that this article is attempting to address, i.e. how to avoid falling into the trap of believing that one can ignore the basic understanding of statistics.

Levels of statistical knowledge

In some schools, medical students are taught basic medical statistics during their pre-clinical training. In our local context, the students in the three medical schools had similar exposure. The disadvantage of the present curriculum is that it tends to stress too much on didactic lectures, rather than their applications. As a result the students find the subject irrelevant to medicine. According to Baker, they found the statistics as peripheral to their main subjects, and do not devote much time to it[3]. Unless the course is structured in such a way that it is more interesting and relevant and continue to be used during clinical years with other courses, this negative attitude will continue as they eventually become doctors. However, the fact remains that the study of medicine is a life long course. Unless the medical practitioners keep abreast with new discoveries and inventions, they may end up being ineffective in delivering good clinical practice to their patients. For most of us, the way we keep abreast with new discoveries are through reading journal articles, and few by attending seminars, conferences and fewer still carry out some research. In all these mode of acquiring new knowledge one inevitably find some numbers, and worse still statistical jargons being mentioned. Take for instance one is likely to encounter such phrases as *'the finding is not statistically significant'*, *'p less than 0.05'*, *'95% C.I.:1.8, 2.7'* and so on. How do we judge a particular article is worthwhile or not if one does not comprehend some of these statistical ideas. To accept blindly what the authors had presented is an act of foolishness on the part of the reader. Can we, in the medical profession afford to continue to ignore the warning "Misuse of statistics is unethical" by Altman[2]?

Some, if not majority of us feel that doing a research does not take much in the way of critically designing a study in order to best answer the study objectives. Also, not much thought is usually given to what sort of statistical analysis that ought to be used. These and other considerations are vital to a good research. often one hears out medical colleagues saying why bother to do all these when there are statistical programmes to do the job. This attitude should be overcome by a sound appreciation of what good research entails.

Continuing statistical education

There is so much that can be done to encourage doctors to change the above perceptions. The essential ingredient for the paradigm shift towards viewing medicine more critically is through their own desire to understand the importance of being able to be critical. Towards this end the stategies of both short and long-term planning is desirable. The related institutions and medical associations could participate in synergising towards this goal.

In our local context, doctors themselves should be encouraged to refresh their basic statistical knowledge by reading one of the many fine textbooks on Medical Statistics. The last decade saw the mushrooming of such textbooks specifically catered for the medical undergraduates and doctors[4,5,6]. These textbooks are well illustrated with lots of medical examples, in contrast to some old textbooks. Another avenue of inculcating proper attitude towards numerical contents of publication is actively participating in journal clubs. Having a moderator who is critical would encourage the members to explore the validity of the findings being reported. And if this goes on, doctors will initiate their own learning.

Statisticians would strongly encourage doctors who intent to embark on a research to seek statistical help from someone who could give advice on designing a research project. This fact is well known to all of us. However, for some reason or other doctors tend to take lightly this advice. But whatever the reasons for such practices, doctors need to seriously rethink about this important advice. Many a time doctors will end up bringing a raw data set to someone and ask if he could run the analysis. Worse still, the objectives of the study were not clear to the 'data analyst' and if he is an obliging person he would try earnestly to help. Some of the time the data may not be salvagable. If pursued with no purpose this may lead to data massaging that would ultimately result in useless information. The ideal situation would be that the investigator had planned his study with some help in the design and type of statistical analysis that would be required. In the eventuality that he further seeks advice, at least he is better equip to discuss the problem more meaningfully.

Statistical package – A tool for abuse?

As mentioned earlier, the statistical packages have to a large extent spoilt the ability of doctors to think critically. As long as he knows which keyboard button to press, he feels confident the programme would come out with a p value to **support** rather than to **illuminate** his study findings. The notion that a statistical programme "gives you an enormous number of canned procedures, so you don't actually need to learn statistics to use it successfully[7]" is to encourage potential users to abuse statistics wantonly. Take the following example where an investigator had used a statistical programme SPSS/PC+[8] to test his data - whether the mean intelligence quotient (IQ) in the two groups of 4-year-old children are the same? As shown by the statistical output (Figure 1), the author was given two options, i.e. two *t*-tests results with *p*-value of 0.016 and 0.122. The question is which of the two *p*-values should the investigator use to interpret his data? Or is there any other procedure, or alternative statistical test that one can use when certain assumptions about the *t*-tests do not seem to hold? Unless the basic assumptions required of the test is known to the investigator, he may be tempted to choose p=0.016 because of its significant result. *('Type 1 error of the second kind!')*

To handle the data in the above manner is not to be encouraged. Statistical packages are most useful and efficient if we know how to use it wisely. This can only be achieved by a strong sense of understanding of some of the underlying principles of statistics.

The investigator wishes to investigate whether there is any difference in 4-year-old children intelligence quotient (IQ) between those born to mothers who suffered postnatal depression and those without. Below is the SPSS statistical output:

t-tests/group depress(1,2)/variables iq.
Independent samples of DEPRESS

Group 1: NOT DEPRESS = 1
Group 2: DEPRESS = 2

t-test for: IQ Intelligence Quotient

	Number of cases	Mean	Standard deviation	Standard error
Group 1	79	112.7848	14.335	1.613
Group 2	15	101.0667	27.004	6.972

Pooled	Variance	Estimate	Separate	Variance	Estimate
t Value	Degrees Freedom	2-Tail Prob.	*t* Value	Degrees Freedom	2-Tail Prob.
2.46	92	0.016	1.64	15.53	0.122

Fig. 1: SPSS programme to calculate 2-sample *t*-test, using observations on IQ in 4-year-old children born to mothers with and without postnatal depression

The future

Nobody should expect improvement in the better understanding of statistics among doctors to come swiftly. Patience is required and with continuing emphasis on the relevance of statistics taught at undergraduate training one would eventually hit the nail right on its head. Nevertheless, one should pursue to convince todays doctors at every opportunity possible. Besides what has been said earlier, there are other ways doctors can be 'forced' to think statistically. Scientific meetings can be one of it - paper presentation should not be limited only to study outcomes, but should also encourage others to present papers concerning methodological as well as statistical issues. In this manner one may generate interest among participants to think. As Albert has said – "One of the important skills a physician should have is the ability to critically analyse original contributions to the medical literature[9]".

References

1. Altman DG. Statistics in medical jounals: Developments in the 1980s. Statistics in Medicine. 1991;10 : 1897-913.
2. Altman DG. Statistics and ethics in medical research. Misuse of statistics is unethical. Br Med J 1980;281 : 1182-4.
3. Baker MR. A student's view on medical statistics. Newsletter. Teaching of statistics in the health sciences. American Statistical Association. July 1989.
4. An Introduction to Medical Statistics. Bland M. Oxford University Press. Oxford. 1987.
5. Medical Statistics. A Commonsense Approach. Campbell MJ, Machin D. John Wiley & Sons. Chichester. 1990.
6. Basic & Clinical Biostatistics. Dawson-Saunders B, Trapp RG. Appleton & Lange. Norwalk. 1994.
7. Pertzold J. Turning data into information. PC Magazine. 1985 November 12 : 155-60.
8. SPSS/PC+ 4.0 for IBM PC/XT/AT and PS/2. SPSS Inc Chicago 1990.
9. Albert DA. Deciding whether the conclusions of studies are justified: a review. Med Decision Making. 1981;1 : 265-75.

MCQS For Medical Statistics : Understanding of Basic Concepts

1. 79 4-year-old children as 'IQ 112.8±14.3' (mean ± standard deviation)
   a) It is 95% certain that the true mean lies within the interval 84.2 - 141.4
   b) Most of the infants had IQ of 112.8, the remainder were having IQ between 98.5 and 127.1
   c) Approximately 2.5% (of 4-year-old children) had IQ of less than 98.5
   d) Approximately 2.5% (of 4-year-old children) had IQ of more than 141.1
   e) Approximately 95% (of 4-year-old children) had IQ between 84.2 and 141.4

2. In the assumption of *t*-test for 2-samples:
   a) The two groups are independent
   b) The number (of 4-year-old children) in each group need not be the same
   c) Either one of the groups IQ must be normally distributed
   d) The variances in the two groups are equal
   e) One can ignore all the assumptions if sample size is very large

3. The observed difference between the IQs means (112.8 and 101.1) can be possibly explained by
   a) biases in the study design
   b) chance
   c) confounding
   d) true difference
   e) few large IQ values (outliers) in group 1

4. *p*-value = 0.016
   a) The probability of the difference in the means IQ (112.8-101.1) being due to chance if null hypothesis is false
   b) The probability of the difference in the means IQ (112.8-101.1) being due to chance if null hypothesis is true
   c) The probability of the differences in the biases during the selection of the sample
   d) The difference in the mean IQ (112.8-101.1) is large enough to be ascribe to chance if the null hypothesis is true
   e) There is only a small probability of obtaining the difference in the means IQ (112.8-101.1) if the depressed mother does reduce the 4-year-old children IQ

5. The difference in the two standard deviations IQ (27.0-14.3) were found to be statistically significant
   a) *t*-test is not a valid test for this data
   b) One may consider using non-parametric tests such as Wilcoxon rank sum test
   c) t-test can still be used provided the sample size is equal in the two groups
   d) One should not pay too much attention to the unequal standard deviations
   e) Based on the above information, it is not possible to interpret the study finding