THE VALUE OF EPIDEMIALOGICAL RESEARCH LOOKING BEYOND THE P-VALUE

SS Hirve, CS Yajnik

Among the sciences, epidemiology is an infant. AIthough some excellent epidemiological studies were conducted before the 20th century, a systematized body of epidemiological principles (Modern Epidemiology) has begun to develop only in the last few decades. These principles have evolved out of a flurry of epidemiological activity covering a wide range of health problems. Some disagreement does exist in relation to basic concepts, including the definition of epidemiology itself.

Epidemiology in Practice

In day to day clinical practice, we are often faced with difficult cases. When making clinical choices we rely on standard teachings. Our past experiences, are sometimes influenced by the opinion of colleagues, by scientific reports and many a times by intuition. All these could be misleading. Clinical reports often fail to compare the abnormal with the normal. Comparisons are crucial in determining whether a treatment improves the course of a disease. Also, many of the clinical questions cannot be answered experimentally for ethical or logistic reasons. Such questions need to be addressed through observational studies, comparing those who are 'naturally' exposed with those not exposed. The science of epidemiology comes of our help in such situations.

Answering the Research Question

Epidemiology is based on two fundamental assumptions: 1) That human disease does not occur at random, and 2) That human disease has casual factors that can be identified through systematic investigation. In his investigations an epidemiologist follows a standard approach, in which he first defines the *disease frequency*, then studies the *disease distribution* in that population and subsequently undertakes a study of *zdisease determinants* with a view to form a strategy for prevention. It is said that anyone who can observe, think logically and make simple calculations can be an epidemiologist. In fact, most doctors behave as one in everyday practice. Unfortunately, our medical curriculum stresses to much on treatment rather than prevention and therefore, epidemiology has become a backyard science. An average medical graduate thus feels shy of epidemiological methods This has contributed to the poor quality of medical research in India and a blind dependence on Western literature which many a times may be quite inappropriate for our problems.

Various epidemiological designs can be used to address the research hypothesis. The design depends upon the problem to be investigated. Many studies are doomed to failure because they fail to appreciate the scope of the study and use inappropriate methodology. Rarity of the condition, the capabilities of the researcher, facilities available and the monetary resources are important considerations.

Generally, though not necessarily, a research process begins with case reports and case series, followed by descriptive epidemiological assessment (prevalence of the disease in different population groups and different geographic locations). Next, case control and cohort studies are designed to identify risk factors and prognostic factors. Lastly, trials (clinical and community) for individuals and groups may be done to assess therapeutic and preventive measures. Although each design has an appropriate role, it is very necessary for clinicians to realize the limitations of each design. The success of any epidemiological research depends on how clearly defined the research question is. No matter how good the design, at best it can only answer the question posed. Different designs are discussed in more detail in a later article.

Errors in Epidemiological Research

The value of any epidemiological study will depend on how well it deals with issues of error or uncertainly in studying associations between exposure and outcome. This is because the association need not be casual. Two relevant terms in epidemiological jargon are precision which refers to reduction of random error, and validity which means a lack of systematic error. Precision is usually achieved by increasing the sample size and validity by improving the representativeness of the sample. Following issues need to be critically addressed to meet these two requirements:

1) chance
2) bias
3) confounding

Chance

The validity of a statistical association is evaluated by the omnipresent 'p-value' which is really the probablity that an effect at least as extreme as that observed in a study could be due to chance alone, given that there is truly no relationship between the exposure and outcome. "Significance" testing is

From Diabetes Unii, K.E.M. Hospital and Research Centre, Pune 411011.

frequently misinterpreted to imply biomedical importance. The preoccupation of an average reader to look for the magical figure of "p<0.05" diverts the attention away from the true biomedical inference to focus on the statistical inference.

Bias

Bias is a systematic deviation from the truth that distorts the research findings. More than fifty biases have been defined. The common ones are : selection bias, information bias, measurement bias, follow up bias, analysis bias, interpretation bias, publication bias and confounding. Certain types of bias are inherent and some types are more prone in certain study designs and one needs to take special care to reduce this error. The key to reducing bias is to identify its potential sources and reduce them at the study design stage itself e.g. *randomization* of study subjects may reduce selection bias in an intervention study design. *Blinding* of subjects, researchers and even statisticians can reduce measurement bias, information bias, bias due to unequal follow up and analysis bias. Meausrement bias can be reduced by use of repeat measures, using multiple sources of information and using measures that are as objective as possible. Bias can be random or systematic in nature. A random (non-differential) bias will distort the association towards the null hypothesis while systematic (differential) bias may distort the association in either direction. However, some amount of bias will always remain in the most perfect study. The key issue should be for the researchers and the readers to be able to assess the importance of the effect of the bias on the study findings.

Confounding

The concept of confounding is of central importance in modern epidemiology. At the simplest, confounding may be considered as a mixing of effects i.e. the exposure effect on an outcome is distorted because of some third extraneous factor. This may lead to over - or underestimation of the effect or even to the reversal of the effect, a phenomenon which is known as the '*Simpson's parado*'. Confounding can occur even if the exposure factor has zero effect on the outcome. For a factor to be confounding, it should have an independent effect on the outcome and also be associated with the exposure factor. The confounding factor should also not be just an intermediate step in the path between the exposure and outcome. Common examples. of confounders are age, sex and social class.

Confounding can be prevented at the design stage itself or can be corrected at the analysis stage. Three methods are commonly used to prevent confounding in the design of epidemiological studies. *Randomization* where subjects are randomly assigned to exposure categories is applicable only to experimental studies. *Restriction* (restricting the admissibility criteria for subjects) prevents confounding by not allowing the confounding factor to vary. Matching ensures identical distribution of the confounding factor in the groups to be compared, however, it is rarely used since it is too costly.

Control of confounding in data analysis is achieved either by *stratificafion* or by *multivariate* analysis.

Cause and Effect: the Unproven Relationship

Epidemiology like most modern sciences relies on statistical principles and applications in data analysis. Statistics however serves only as a tool for achieving the scientific objectives. Too often, statistical procedures are employed in epidemiological studies with little regard to the epidemiological objectives.

A causal association is one where a change in frequency or quality of an exposure results in a corresponding change in the outcome of interest. Any study which shows an association may at best suggest (but will never prove) a causal relationship between an exposure and an outcome. Causality is indirectly suggested by the *strength of the association*, the presence or absence of a *dose-response* relationship, showing the *temporal sequence* of the outcome following the exposure, *consistency* of results with other investigations in different settings, the *specificity* of the findings, whether the outcome is *reversed* on reducing the exposure and finally by the bio-medical *credibility* of the association. Epidemiology and biostatistics at most establish associations and suggest causation but rarely if at all, will establish causality.

Integrated Research - The Need of the Future

Epidemiology is said to be the simplest and most direct way of studying the causes of a disease in man. It only requires an ability to count, to think logically and to think imaginatively to be an epidemiologist!

However, given the multidimensionality and complexity of human disease, epidemiological research is increasingly becoming a multi-disciplinary team work. Only in this way we can bring together the epidemiological skills, clinical experience, biological understanding, statistical expertise and many other special skills to understand the health and disease of the human beings. Even though epidemiological research is becoming more complex, the core of the subject remains simple and a good study should be able to describe in such a way that all those who are interested in the cause of the disease can follow the arguments and decide for themselves the validity of the conclusions.