

# How to Critically Appraise a Medical Article

Pages with reference to book, From 77 To 79

Gohar Wajid ( National Health Research Complex, Shaikh Zayed Hospital, Lahore. )

Medical research, like many other disciplines of life is more concerned with 'quality' rather than 'quantity'. On the basis of qualitative assessment, probably no article can be declared as a 'perfect article' as every article has its strengths and weaknesses. Qualitative assessment of medical articles is a complex and technical issue and requires, atleast basic epidemiological and biostatistical knowledge. The purpose of reading medical articles is to get as close as possible to "the truth" upon which one bases a decision such as, how to treat a patient or what prevention action works in a community? As no study is ever perfect, we nina riskofrejectingallthe articles if our expectations are too high<sup>1</sup>. The best approach is thus to rank the articles by their quality. This paper describes a common approach to evaluate the strengths and weaknesses of medical articles and thus to assess their quality. Issues related to study objectives and study design. For any study, it is important to note, how the author has stated the introduction of the study. Are the definitions of important terms stated unambiguously, supported by appropriate references? One of the purposes of literature review is to provide justification for performing the study. For example, it may not be appropriate to conduct a study to prove that smoking causes lung cancer, as performing such a study nowadays, is equivalent to re-inventing the wheel and may be unethical as well. Every article should give a sound justification for conducting research on the specific topic. This justification is primarily supported by literature review, stating references from other medical articles, critically analyzing the strengths and weaknesses of these articles, thus identifying the gaps in medical research and the need of present study. In a good article, the author should clearly mention the objectives of the study. Where-ever applicable, goals and objectives of the study must be specified by time, person, place and amount<sup>2</sup>. The objectives often have sound relationship with the design of the study and should give an idea of the type of study being conducted. The objective of a descriptive study may be "to describe or to assess" and that of analytical study may be "to analyze". A good study should reflect the research questions being addressed by the author. For example, two different research questions may be:

- 1) What is the magnitude (frequency) of asthma in children under 12 years in Lahore?**
- 2) Is smoking a cause of lung cancer?**

Most of the time, the research questions in medicine are related to the impact of an intervention, causality, evaluation of a diagnostic test or determining the magnitude of a problem. In a good study the reader should be able to identify the study type used by the author to address the research question. A cross-sectional study is a suitable study type to answer the first research question and a case control (or may be a cohort) for the second research question, mentioned above. The reader should be able to appreciate the suitability of study type for addressing a particular research question. One of the purposes of assessing the study type is to see whether the objectives are achievable through the adopted study type and what could be the best possible study type to achieve the stated objectives. Issues of causality cannot be discussed by a cross-sectional study design, thus a cross-sectional study will not provide answer to any research question related to causality. A study can be declared as the best, if it has least bias in it. Three general types of biases threatening the internal validity of any study are selection bias, information bias and confounding.

## **Issues related to selection bias**

For any study, it is important to identify how different subjects are selected from the study base (population of interest). Poorly defined exclusion and inclusion criteria, poor study base, non-random selection, refusal to participate, loss to follow-up and even missing data may all induce bias in the study. This type of bias is called selection bias. Selection bias threatens the internal validity of the

study. It is defined as the distortion of the effect measured, resulting from procedures used to select subjects that lead to an effect estimate among subjects included in the study different from the estimate obtained from the entire population theoretically targeted for the study and internal validity is defined as the validity of the inferences drawn as they pertain to actual subjects in the study<sup>3</sup>. Improper assigning of subjects to different groups and re-grouping of the subjects to different groups and re-grouping of the subjects at any stage during the study may also introduce selection bias. Even at analysis stage, it is appropriate to avoid re-grouping and if desired, should be done with great caution. A good study should also mention whether all the cases divided in different groups at design stage of the study, also reached the analysis stage or there were any drop-outs in between. These drop-outs may be in the form of death of the subject and non-participation of the subjects due to any reason. The true results may become inflated or deflated because of missing persons. If there are any drop-outs, the effects of these on the study should be discussed.

#### **Measurement issues leading to measurement bias**

All the study factors and outcome factors in the study should be mentioned clearly and the method of measurement of each factor should also be stated precisely. Any inaccuracy in the measurement of study factor or outcome factor (or even confounders) may lead to a type of bias, called measurement or information bias. When a continuous variable (such as height, blood pressure) is involved, the associated error is called 'measurement error' and when a categorical or discrete variable (such as pulse, White Blood Cell count) is involved, the related error is called 'misclassification'. For example, different blood pressure apparatuses may give slightly different readings for the same subjects (called instrument error) or the doctor may have hearing deficit (observer error) or the subject may be little anxious at the time of examination, thus giving slightly different reading than actual (subject error). Misclassification is defined as erroneous classification of an individual, a value or attribute into a category other than that to which it should be assigned. The probability of misclassification may be the same in all study groups (non-differential misclassification) or may vary between groups (differential misclassification)<sup>4</sup>. These errors may lead to measurement bias and thus distortion of results from true value.

- Differential misclassification occurs when the misclassification probabilities (sensitivity and specificity) differ between study groups, that is, groups classified by exposure or outcome. Case control studies are particularly vulnerable to this type of error because of possible intensive interviewing of the cases than controls or better re-call of study factor by the cases. For any such study, it is often difficult to know whether bias exists or not and what the magnitude and direction of bias may be. All we can do is to examine whether the study methods have been optimized to avoid bias and to judge how likely it is that bias may still occur. Differential misclassification can be reduced by improving comparability of measurement techniques between study groups. For example, using the same interviewer for all the subjects in the study. In any study it should be noted whether it tells about the relevant outcomes (content validity). In descriptive studies, lack of measurement validity (different interviewers, improperly calibrated instruments) may lead to over or under estimation of the health problem being measured. When reading such studies, ensure that the methods of measurement of different variables have been stated properly. In RCT, Cohort or Case Control studies, find out whether the author has mentioned the action taken to avoid differential misclassification.

#### **Issues related to confounders**

Another important factor to be observed in an article is possible role of potential confounders. This is especially important for cohort studies. It has also been observed that randomization may not totally solve the problem<sup>5</sup>. Confounding variable or confounder is defined as a variable that can cause or prevent the outcome of interest, is not an intermediate variable and is not associated with the factor under investigation<sup>6</sup>. Retrospective studies, prospective or cohort studies and experimental studies must take into account confounding variables. Confounding variables can be recognized and prevented

at the beginning of studies (design stage) by the process of pairing or matching, or they may be taken into account or controlled for after the data have been collected at analysis stage by using statistical techniques, as regression analysis. For example, if we look for causal relationship between cigarette smoking and cardiovascular diseases, at the same time we have to look for other factors likely to affect this causal relationship, such as alcohol consumption, high cholesterol level and high blood pressure.

### **Issues related to statistics**

One should also look for various statistical procedures applied for expressing the results. Is the sample size enough to detect a clinically or socially significant result? Are all the hypotheses tested by applying appropriate statistical techniques? Has the author mentioned the values often or  $\chi^2$ , degree of freedom, standard error, level of probability, confidence intervals, different measures of frequency or association, as appropriate. It is also important to note whether the author has correctly interpreted the results. Can we draw conclusive results from the study, if not, why? Has the author mentioned the reasons for inconclusive results? It may not be important whether the results of the study are statistically or clinically significant or even sometimes, whether the study is able to produce conclusive results, what is probably more important in my opinion is whether the study has achieved the stated objectives or not. Similarly the generalizability of the study may not be as important as its validity.

### **References**

1. Department of Public Health, University of Sydney, How to critically appraise the literature, Sydney, University of Sydney Pub., 1991, p. 86.
2. Hawe, P., Degeling, D. and Hall, 3. Evaluating health promotion, MacLennan and Petty, 1990, p.46.
3. Rothman, K.J. Modern epidemiology, Boston, Little Brown and Company, 1986, pp. 82-83.
4. Last, J.M. A dictionary of epidemiology. N. York, Oxford University Press, 1988, pp.29,62.
5. Altman, D.O. Comparability of randomized groups. Statistician, 1985;34: 125-36.
6. Riegelman, R.K. and Povar, G.J. Putting prevention into practice, problem solving into clinical prevention. Boston, Little Brown and Company, 1988, p.22.