

Perspective on Variables in Medical Research

B. I. Avan (The human Development Programme, The Aga Khan University, Karachi.)
F. White (Department of Community Health Sciences, The Aga Khan University, Karachi.)

Abstract

Research concepts cannot be materialized until and unless study variables are carefully selected and clearly defined. However, emphasis on variables is not limited to data collection. It should be envisioned from the start that the contents of the variables recorded are congruent with the statistical analysis suggested. These two processes require a comprehensive understanding of variables in epidemiology and biostatistics disciplines. In this article classifications are discussed keeping in view their utility in medical research (JPMA 51: 94, 2001).

During fellowship exams, the candidate presented the history of a patient lying on a bed, ' he was shocked when the patient narrated a totally different version of history in front of' the examiner Even his vital signs behaved differently as if they had never been measured before....

Every biological measurement varies from person to person and within the same person at different times. It would be virtually an "out of this world" experience if we found a group of people of exactly the same height, same weight, same blood pressure, same heart rate, same speech and thoughts.... For a medical practitioner, variation is a day-to-day experience. Even the provision of excellent diagnosis and medication cannot assure 100% effective cure in a patient. There are numerous reasons, which make that person and disease state different from others. Researchers always struggle to identify the factors that produce such variations. They strive to identify and measure the change from one value to another. Their eventual emphasis is to capture the dynamic nature of the attribute in order to achieve the ultimate aim of the science, i.e. to make the phenomenon or event more predictable. Variation can be assessed at various levels, it can be intraindividual and inter-individual.

The secretory rates of adrenocorticotrophic hormone and cortisol are all high in the early morning but low in the late evening¹.

Sometimes, variability is studied specifically to differentiate the abnormal from the norm, pathology from physiology and disease from health.

Blood pressure greater than 220/150 mmHg is an immediate threat to life and emergency treatment should be given for acute hypertensive crises².

Sometimes, it is studied in order to control it and so that the underlying mechanism can be investigated Pre-school children are stratified into malnourished and well nourished groups to assess their health seeking and hygiene behaviors³

Moral: Science exists because things vary with some regularity. Edmond A. Murphy.

Variables are the specific properties that have the ability to take different values. Often, the discussion of variability and variable encompasses the topic of variance that is a distinct statistical concept.

Variance refers to deviation of the individual values from the mean of the sample.

Variables can be classified differently according to their utility and the vision of the scientific discipline. For medical research, an understanding of variables is important from both epidemiological and biostatistical perspectives.

In epidemiology, variables are considered in terms of description or presentation, and relationships. The allocation of specific variables into the specific classes is relative. It depends upon the level of knowledge, study questions and methodology.

Demographic variables describe the characteristics of the study sample.

To study the problems of children of' divorce, the sample consisted of 356 children (176 girls and /80 boys) whose mothers were divorced. The mean age of the children was 10.6 years. The majority of the mothers (86%) were white⁴.

Such information is used to assess that the proper study sample has been selected to answer the study questions and that it provides a scenario that allows study findings to be generalized.

Exploratory variables describe the characteristics of the phenomenon or process under investigation. Identify the disease types thought to be treated appropriately by Western medicine by doctors trained professionally in Western medicine and Traditional Chinese medicine⁵.

The variables at this stage are not meant to examine the cause and effect relationship. They just explore what are the possible components of the phenomenon. They may also form the basis for hypothesis generation and for comparative studies.

Dependent variable (syn: Effect, criterion, criterion measure, outcome, output variable) is a response that the researcher wanted to predict.

Independent variable (syn: treatment, experimental. predictor, input, exposure. explanatory variable) is a stimulus or activity that is identified or manipulated to predict the dependent variable. The assumptions regarding the association that the independent precedes the dependent variable and there is a cause and effect relationship between the two.

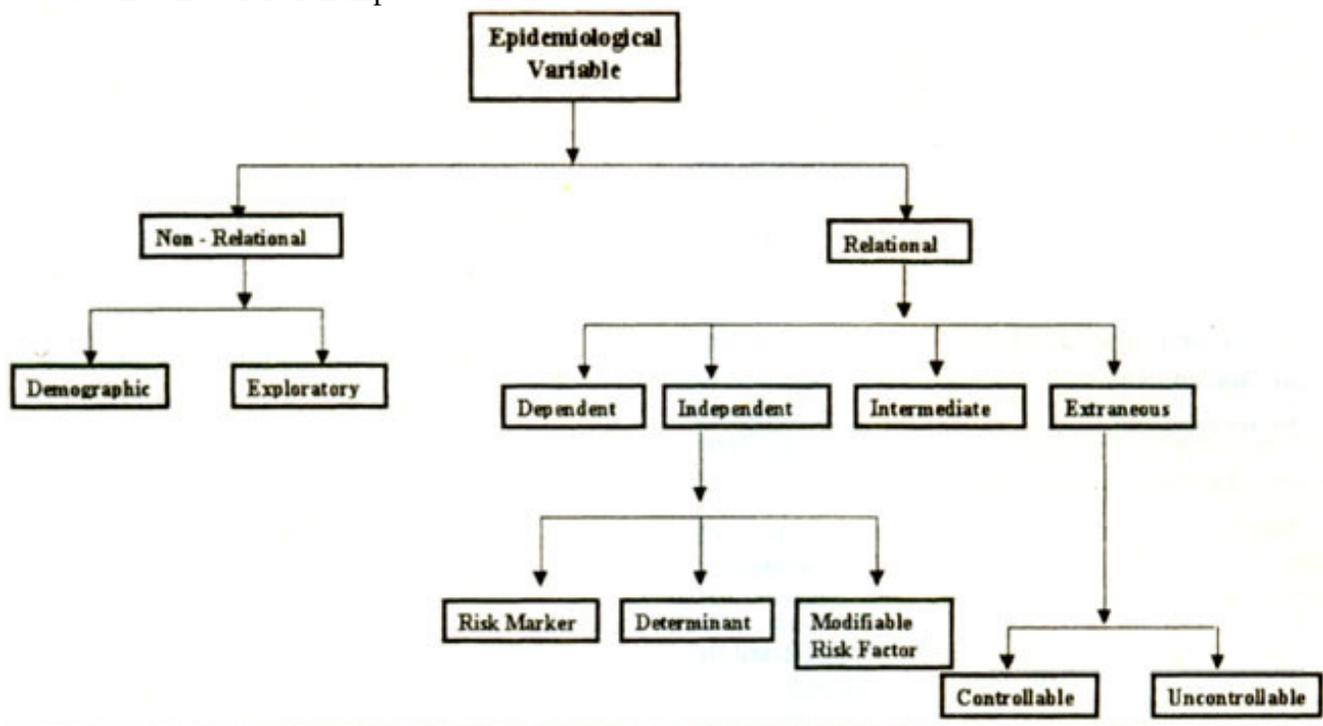


Figure1. Classification of variables in Epidemiology.

A BMI either high or low, was associated with reduced probability of' achieving pregnancy in women receiving reproduction treatment⁶.

In epidemiology the independent variable has numerous levels.

Passenger variable refers to a spurious association in which passenger variable varies systematically with the dependent variable. It is explained by the third variable that is associated with both passenger and dependent variables.

The association of neonatal tetanus with topical use of clarified butter on the umbilical wound becomes nonsignificant, when the association was controlled to exclude the dung fuel used to warm the clarified butter for topical use Risk factor is an independent variable determined by the evidence provided by the

epidemiological studies. This generic term is used for the concepts of risk marker, determinant (factor) and modifiable risk factor.

Risk marker is an independent variable associated with an increased likelihood of happening of specified outcome variable.

Alpha fetoprotein is a tumor marker associated with hepatocellular carcinoma⁸.

Determinant (factor) is an independent variable that increases the likelihood of the happening of a specified outcome. When the factor can be manipulated to prevent the happening of a specified outcome, it is called modifiable risk factor.

Asbestos is the risk factor for lung cancer among shipboard personnel⁹.

Intermediate variables are those that occur in causal pathway from an independent to a dependent variable. Causality is a chain reaction. One outcome forms the basis of the independent variable for another outcome. eg. poor working environment leads to job dissatisfaction, which in turn leads to absenteeism. Now if the question under investigation is the effect of working environment on absenteeism, then one of the potential effects of poor working environment would be mediated by job dissatisfaction

Extraneous variable is a variable that has a potential to distort the relationship between dependent and independent variables.

Controlled extraneous variables are recognized before the study is initiated and are controlled in the design and selection criteria.

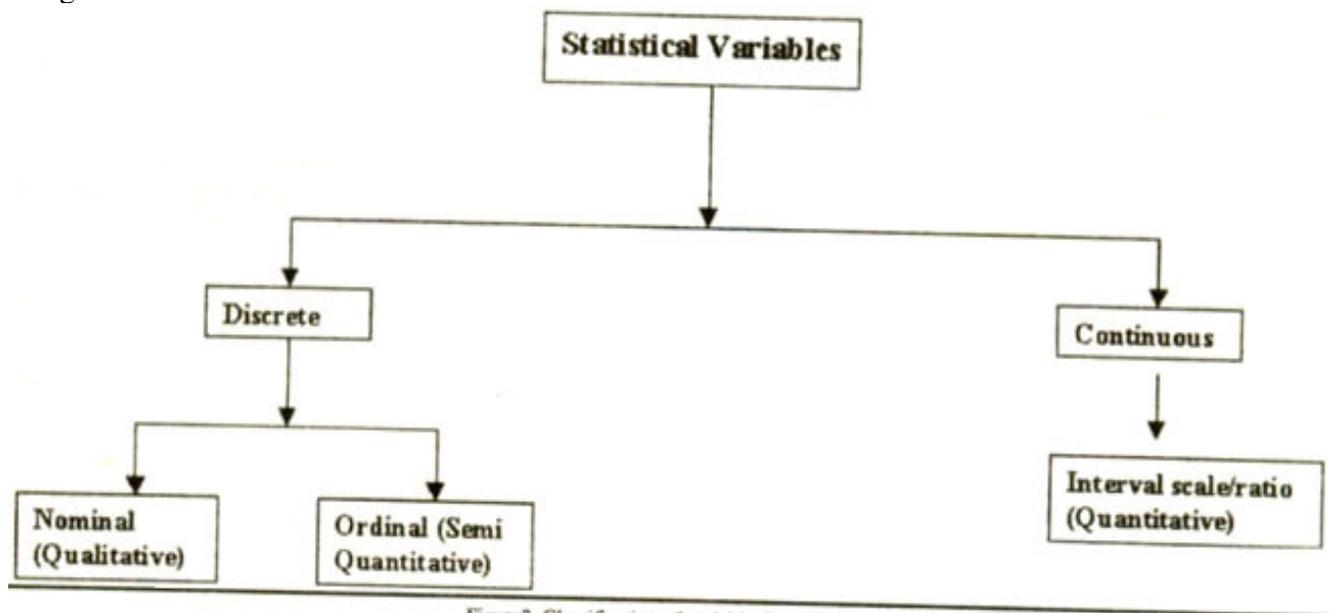


Figure2. Classification of variables in statistics.

In order to study the home care of malaria infected children in the rural area of Guinea, study subjects of less than 5 years of age were selected and had to have lived in the study area for at least six months¹⁰.

Uncontrolled extraneous variables (confounding) are recognized before the study is initiated or, sometimes, even if recognized cannot be controlled in the design and selection phase. Usually an attempt is made to assess and adjust them through sophisticated statistical tools.

The presence of a congenital anomaly increased infant mortality 9 fold (95% CI = 7.3-11.1) for black infants and 18 fold (95% CI = 16- 17.8) for white infants¹¹.

Environmental variables are those extraneous variables that form the set up in which non- experimental studies are conducted. They are not intended to be controlled as the study may lose its context.

Data from 27 European countries show that poor environment during infancy and childhood, which is

associated with high infant mortality, may explain some of the similarities in the description of epidemiology of stroke and stomach cancer¹².

In statistics, variables are considered quantitatively and depend upon study question and methodology, level of knowledge, availability of measurement tools and resources.

	Quality Variation	Ranking Variation	Actual Quantity	Statistical Procedures Applicable
Nominal	Maximum	Nil	Nil	Minimum
Ordinal	Moderate	Moderate	Nil	Minimum
Interval Rate	Minimum	Maximum	Maximum	Moderate
Interval Ratio	Minimum	Maximum	Maximum	Maximum

Figure3. Characteristics of variables in statistics.

Discrete variables assume that fixed or limited values are arranged into contrasting groups.

They have two main sub classifications:

Nominal variables: the observations having similar characteristics or quality are grouped together into categories i.e. contrasting categories with descriptive titles.

Race, Religion, Marital status, Sex etc.

The utility of nominal variables seems minimal, but they are very useful in qualitative research. Furthermore, in quantitative research, it is neither always feasible nor required to quantify every encountered variable. Sometimes it is the only way left until the scientific community is able to reach a consensus to measure the quality of a variable.

Ordinal variables: the basis of categorization is not only the similarity in characteristics but also the similarity in quantity. Furthermore, categories are arranged and ordered only in accordance with their quantity and not according to the standard measuring units. A word of caution here is that neither is there restriction of the same quantity in each category, nor is the inter-category distance necessarily equal.

The long distance learning in radiology was assessed on the likeri scale, by ranking quality of session as excellent = 4, good = 3, fair = 2, unsatisfactory = 1, poor If the discrete variable, either nominal or ordinal, has only two mutually exclusive categories, it may be also labeled as dichotomous variable.

Male/female, younger/ older etc.

Interval variables: The measurements are based on standard measurement units and intervals between the intervals are equidistant. Measurements are possible both above and below zero point. Whenever the “0” point is arbitrarily decided, it is called interval scale. Whenever the “0” point is real it is called interval ratio. The difference is usually cleared by a classical example of weight and temperature.

Ordinal: one of the age classification for demographic characteristics of the population.

i) Infants and children (0 - 14 years).

ii) Young people and adults (15 - 49 years).

iii) The aged (50 and above).

Interval ratio: actual time considered in years, months, and days etc

Moral: Epidem iological variables for practical purposes may be defined according to a nominal, ordinal and interval/ratio scale.

Simultaneous development of understanding of epiderniological and statistical concepts is indispensable in medical research. Biostatistics plays a very influential role in medicine primarily for the reason of enhancing the validity of clinical measurements as well as the findings of research. While medical professionals try to understand the related epidemiological principles, they often feel uncomfortable with statistics. However both disciplines are important in the critical appraisal of the medical literature and also to properly use quantitative information in practice. On the other hand, novice researchers sometimes have a misconception that data analysis is simply giving a command to

the computer will virtually automatically result in valid and complete statistical analysis. This assumption is both false and potential hazard.

Selection and defining variables are the first steps in medical research where statistics are required to proceed in epidemiological studies.

Moral: Computers are only as knowledgeable and skilled as their operator is.

References

1. (Ivton AC. Textbook of Medical Physiology. Chapter 77, The adrenocortical hormone. 8th edition Philadelphia, WB Saunders Co 1991.
2. Charles E. S. Maier HO. Emergency diagnosis and treatment. Chapter 28. Cardiac emergencies (468 - 485) 4th edition. Connecticut, Appleton and Lange., 1992.
3. Abate O, Kogi-Makau W, Muroki NM. Health seeking and hygiene behaviours predict nutritional status of pre-school children in a slum area of Addis Ababa, Ethiopia. *Ethiop Med. J.* 2000;38:253-65.
4. Sandier Ni. Tein JY. Coping efficacy and psychological problems of children of divorce. *Child Dev:* 1998;71 1099-111.
5. Harmsworth K. Lewith OF Attitudes to traditional Chinese medicine amongst Western trained doctors in the Peoples Republic of China *Soc. Sci. Med.*, 2001;52:149-53
6. Wang JX. Davies M, Norman RJ. Body mass and probability of pregnancy during assisted reproduction treatment: retrospective study, *Br. Med. J.*, 2000 32 1:1320-1.
7. Bennett J, Ma C, Traverso H, et al . Neonatal tetanus associated with topical umbilical ghee: covert role of cow dung. *Int. J. Epidemiol.*, 1999;28:1172-5.
8. Barbara I. B. Pathophysiology : adaptation and alterations in function. Chapter 3, Ncopleia (43 - 63). 4th edition. Lippincott Philadelphia Martin JC, Imbernon E, Goldberg M, Chevalier A. Bonenfant S Occupational risk factors for lung cancer in the French electricity and gas industry: A casecontrol survey nested in a cohort of active employees. *Am. J Epidemiol.* 2000;151 902-12.
9. Last tM Public health and preventive medicine Chapter 17: Asbestos and other fibers (343 -364) 13th edition. Connecticut, Appelton and Lange, 1992.
10. Diallo AB. Serres GD. Home care of malaria of less than 5 years of age in rural areas of the Republic of Guinea. *Bull Wi 10 .* 2001, 79:28 - 32.
11. Malcoe LH, Shaw GM. Iammer EJ. et al. The effect of congenital anomalies on mortality risk in white and black infants *Am. J. Public Health.* 1 999;89:887-92.
12. Leon DA. Smith GD. Past infant mortality predicts current adult mortality from stomach cancer and stroke. *Br Med* 2000;320(7251).F.
13. Tachakra S. Dutton D. Long-Distance Education in Radiology via a Clinical Telemedicine System. *Telemed J. E. Health.*, 2000;6.361-65.
14. Barker DiP, Hall Ai Some important epidemiological variables and attributes, Churchill Livingstone. U.K.. 1991. Chapter 5, pp. 55-80.