Academia Dysphrenia

ISSN 2394 - 2053 (Print)
ISSN 2394 - 2061 (Online)
www.ojpas.org

COMPENDIUM

Open Journal of Psychiatry & Allied Sciences

Epidemiology and biostatistics: fundamentals of research methodology

Abstract

Epidemiology and statistics is an essential modern science whose understanding now forms a base of any research in any form. Epidemiology deals with the distribution and determinants of health-related states, while biostatistics is considered to be important in implementing statistical knowledge into the biomedical sciences. We have tried to capture the basics and essence of epidemiology and biostatistics.

Keywords: Medical Sciences. Sampling. Diagnostic and Therapeutic Materials.

Dhrubajyoti Bhuyan[1], Neha Dua[2], Tejal Kothari[3]

[1]Assistant Professor, Department of Psychiatry, Assam Medical College & Hospital, Dibrugarh, Assam, India, [2]Post Graduate Trainee, Department of Psychiatry, Assam Medical College & Hospital, Dibrugarh, Assam, India, [3]Post Graduate Trainee, Department of Psychiatry, Assam Medical College & Hospital, Dibrugarh, Assam, India

Corresponding author:
Dr. Dhrubajyoti Bhuyan, Assistant Professor, Department of Psychiatry, Assam Medical College & Hospital, Barbari, Dibrugarh-786002, Assam, India. dr.dhrubajyoti@gmail.com

Received: 05 August 2015
Revised: 04 November 2015
Accepted: 05 November 2015
Epub: 11 November 2015

DOI: 10.5958/2394-2061.2015.00022.1

Introduction

Basics of epidemiology and biostatistics are of utmost importance in understanding and formulating a research project. The term epidemiology is derived from the Greek word 'epidemic' (epi=among, and demos=people; logos=study). It is a very old word dating back to the third century BC. According to John M Last, epidemiology is "The study of the distribution and determinants of health related states or events in specified populations and the application of this study to the control of health problems".[1] Clinical epidemiology extends the principles of epidemiology to the critical evaluation of diagnostic and therapeutic modalities in clinical practice.[1]

Statistics comes from Greek word status, meaning 'state' or 'position'. Biostatistics is concerned with the development of statistical theory and methods, and their application to the biomedical sciences.[2]

History

In the earliest times, statistics was used for affairs related to administration of various matters in the country. Then, some insurance companies started using statistics to find out longevity of people in order to fix insurance premium for various age groups. Thereafter, data on births and deaths began to be collected in western countries. John Graunt's 'bills of mortality'[3] and Williams Farr's 'systematic compilation of causes of death' done in Registrar General's Office in England[4-6] were noteworthy studies. They became landmark studies for data collection and vital statistics. Work of Mendel on genetics was the path breaking event in the history of biostatistics. First medical book on biostatistics was by Austin Bradford Hill in 1937. Architect of modern statistics in India is considered to be PC Mahalanobis, who was helped by distinguished scientists like CR Roy, RC Bose, SN Roy, SS Bose, KR Nair, DB Lahiri, and many others. There is a mention of probability in Mahabharata, story about King Bhangasuri and Nala. Arthashastra by Kautilya gave details of data collection on agricultural population and economic census.[7] The first regular census in India was taken in 1881 and others took place at ten-year interval.[1]

Aims and Scopes of Epidemiology

Epidemiology has three main aims, according to the International Epidemiological Association (IEA):[8]

a) to describe the distribution and magnitude of health and disease problems in human populations,
b) to identify aetiological factors in the pathogenesis of disease,

c) to provide data essential to the planning, implementation, and evaluation of services for the prevention, control, and treatment of disease, and to the setting up of priorities among those services.

The knowledge of epidemiology helps us to study the natural history of a disease and to measure disease frequency in terms of magnitude of the problem. It also helps to make community diagnosis, formulate aetiological diagnosis, and identify the risk factors of a disease. It has a role in estimating the individual's risk of a particular disease, identifying syndromes, formulating the plan of action evaluating the health services and making researches.[1]

Uses of biostatistics include documenting natural and medical history of psychiatric diseases, planning clinical trials, evaluating various levels of treatments, providing standard measures of accuracy of clinical procedures, and predicting outcome of common psychiatric disorders. It is also helpful in assessing state of mental health of community, planning and monitoring various mental health programs for specific population groups; thus, eventually assessing their failure or success and promoting mental health legislation.[2]

Steps for Experimental Study

Steps for an experimental study start with asking a question, defining a problem, and its aims and objectives. Following this, an extensive review of existing literature or background research is done and a hypothesis is established or constructed. Hypothesis is tested by doing an experiment. For this, a plan of action is set up after deciding the nature of study. This plan includes defining the population under study, selecting samples from the population defined, ruling out errors recording the data, formulating a work schedule. Finally, the data collected and analysed, and is presented along with the results and conclusion derived from the experimental study.[9]

In statistical inference of observed scientific data, null hypothesis (H0) refers to a general statement or default position that there is no relationship between two measured phenomena. In statistical significance, null hypothesis is generally assumed true until proved otherwise.[7] There are two mentionable approaches of significance testing: Ronald Fischer approach and Jerzy Neyman Egon Pearson Approach. H1 is alternate hypothesis.[7]

Hypothesis testing is not a full-proof of validity of a given statement (or, hypothesis). There are two types of errors that occur in hypothesis testing. Type-I error is said to have occurred when a hypothesis is rejected, but actually it was true. Type-II error is said to have occurred when a hypothesis is accepted, but actually it was false.[7]

Sampling is done when a large population is there for the study. Representative sample resembles/represents the population in terms of characteristics under study. An element/group of elements of the population used for drawing a sample is sampling unit. Sampling frame is the list of sampling units with identification addresses in a population which serves as a base for selecting a sample. Sampling fraction is proportion of sampling size chosen to a population, whereas sampling interval is the inverse of sampling fraction.[2,7,9]

Sampling bias can be present. In a study, as the sample size increases, error chances decrease; but, cost increases. It is very important to find the optimal sample size. Minimum sample size required is calculated on the base of study design and its specific objective along with other possible selective statistical considerations.[2,7,9]

Minimum sample size (N) = $[Z^2 1- aP(1-P)]/d^2$

where a = Level of significance, Z = 1.96 (standard), P = Prevalence rate, d = Level of precision (at 5%, i.e. p=0.05).[10]

Probability sampling methods include simple random sampling, stratified random sampling, systemic random sampling, multistage random sampling, and cluster sampling. Judgement, quota, convenience, snowball, and extensive sampling are non probability sampling methods.[2,7,9]

Simple random sampling is done by assigning a number to each of the units. A table of random numbers is then used to determine which units are to be included in the sample. Systematic random sampling is done by picking every fifth or tenth unit at regular intervals. In stratified random sampling, the sample is deliberately drawn in a systematic way; so that, each portion of the sample represents a corresponding stratum of the universe. The population under study is first divided into homogeneous groups called strata and the sample is drawn from each stratum at random in proportion to its size. Multistage sampling method refers to the sampling procedure carried out in several stages using random sampling technique and is usually employed in large country survey. A cluster is a randomly selected group. Cluster sampling method is used when units of population are natural groups or clusters, such as villages, wards, blocks, factories. In multiphase sampling method, part of information is collected from the whole sample and part from the sub-sample. [2,7,9]

Epidemiological Methods

Epidemiological studies can be classified as observational and experimental. Depending on allocation, experimental studies can be divided into randomised and nonrandomised. In observational studies, if a comparison group is present, it is an analytical study; otherwise, descriptive study. Analytical study can be further divided as ecological cross-sectional study, case-control and cohort study.[7] Ecological study uses

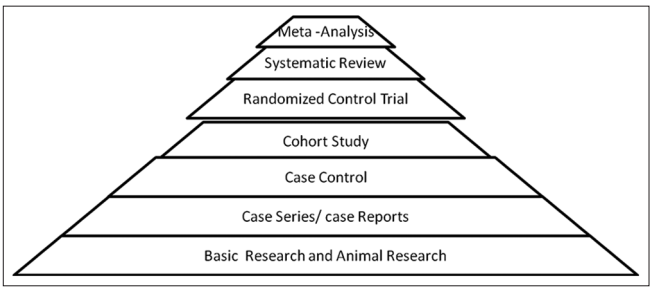

Figure 1: While basic research and animal research is considered the most elementary kind of research, meta-analysis is considered the best and most widely accepted method. Systematic reviews and randomised control trial (RCT) are considered better than cohort study which is above case control study to establish any hypothesis.[11]

population as a unit of study; while individual is the unit of study in the rest of the analytical studies.

Cross-sectional study generates one time snapshot of the situation that exists in a population at a specified time. The exposure under investigation and the disease status are assessed simultaneously. Time of survey is important. Such types of study are more useful to generate rather than test or confirm a hypothesis, which can be done by method of difference. Say if we observe great difference between occurrence of disease between two groups of population, it gives rise to question whether one group is more prone to a disease or not. Another way is by method of agreement. If level of disease is constantly present in some groups, it gives rise to hypothesis that some factors could be common amongst them.[1,7,9]

Cross-sectional study is simple, easy to carry out, lesser cost, lesser effort, provides quick answers, and generating a hypothesis is easy. But, they cannot determine temporal sequence, i.e. when exposure preceded the disease. Moreover, they are insensitive to change in exposure status of study participants. Hence, when a study factor is stable, unlikely to change, a cross-sectional study is the suitable design.

Case-control study is a retrospective study where population under study is observed, examined, investigated, or questioned in order to obtain information about the population with reference to characteristics which occurred previous to the time of survey. In cases, outcome or disease is already present, whereas in controls, outcome or disease is absent. Past exposure is seen and compared; thus, we move from effect to cause. Cases and controls should be similar in all ways, except the disease or variable under study. We ensure this by 'matching'. Analysis is done by odds ratio (OR) which is a way to measure how strongly two variables are associated. If OR is more than one, there is strong association. The OR cannot be negative. It can range from zero to infinity. If odds among exposed is less or OR is less than one, the exposure is said to be a protective factor for the disease.[1,7,9]

The process of making a study group and a comparison group comparable to each other with respect to extraneous factors is known as matching.[12] In individual (paired) matching for each case, one (or more) controls with the relevant characteristics matching the case are chosen. Frequency matching is when controls are selected such that the distribution of the relevant characteristic in the controls is similar to the distribution in the cases. Matching improves the study by removing a strong confounder and study can be done when no sampling frame is available. On matching a characteristic, you create an equal distribution in the cases and controls. Therefore, disadvantage is one cannot examine the association between the matched characteristic and the outcome. Overmatching and under matching can also cause erroneous results.[7]

Cohort is a group of similar subjects, free from disease. We collect a cohort and then look forward to a group of development of an event of interest at a later date amongst us. We observe that amongst exposed people how many developed disease and when, to compare the rates of incidence in different exposure strata. Cohort study is a prospective study.

Examples of types of cohorts are community cohort, exposure cohort, birth cohort, occupational cohort, diagnosed cohort, treated cohort, etc. In dynamic cohort, different subjects enter at different points in time and in fixed cohort, all members of cohort are enrolled in one go. Analysis of cohort helps to establish relative risk, incidence of disease, attributable risk.[2,7] Incidence rates among exposed are obtained by dividing the total exposed who developed disease by the total number of people exposed to risk factor. Incidence rates among not exposed are obtained by dividing the number of people among non-exposed who developed the disease by the total number of people who were not exposed to risk factor. Relative risk (RR) is incidence of disease among exposed divided by incidence of disease among not exposed. Attributable risk (AR) is the percentage of (incidence of disease among exposed minus incidence of disease among not exposed) divided by incidence rate among exposed.[1,7,9]

In experimental studies, a trial is a planned experiment designed to assess the effects of a new treatment/intervention by comparing the outcome of interest in a group exposed to it with those observed in a comparable group receiving a control treatment or intervention. The group receiving the trial is the experimental group and other is the control group. Types are therapeutic, controlled, randomised, and non-randomised.

Randomised control trial (RCT) is an epidemiological experiment. The basic steps include drawing a protocol, selecting reference and experimental populations, then randomisation followed by manipulation or intervention. Follow-up of experimental and control group subjects is done at definite intervals of time, in a standard manner, with equal intensity, under the same given circumstances, in the same time frame till final assessment of outcome. Attrition is the losses to follow-up due to inevitable factors. Finally, the outcome is assessed.[1]

Errors of assessment of outcome due to human element may result in bias. They can be from three sources: Subject/participant variation, observer bias, evaluation bias by the investigator. In order to reduce bias, blinding is adopted. Blinding can be single, double, or triple. Ideally, triple blinding should be used; but, double blinding is the most frequently used method.[1]

Common types of RCTs are clinical trials, preventive trials, risk factor trials, cessation experiments, trial of aetiological agents, and evaluation of health services.[1] Clinical trials are designed to assess the effect of a new treatment/intervention by comparing the outcome of interest in a group exposed to it with those observed in a comparable group receiving a control treatment/intervention. There are four phases of clinical trials. Phase one aims to identify tolerable dose and study pharmacokinetics, and is usually performed on healthy human volunteers. Phase two is a small scale investigation, exploratory study to select drugs/dosages. It is performed on patients with condition under study. Full scale evaluation involving randomisation is done in phase three and a substantial number of patients are involved. Phase four trials are done for post marketing surveillance.[7]

Meta-analysis is simply put as conducting research over previous research. We contrast and combine results from different studies hoping to identify patterns, sources of disagreement, and relationships between the studies. Data collected by the study can be primary and secondary. If the budget is sufficiently large, the researcher collects the primary data for the purpose of study, e.g. through field surveys and laboratory experiments. Data extracted from records maintained by other agencies for some other purposes is secondary data, e.g. in psychiatric research, major data comes from records of government mental hospitals, private psychiatry nursing homes, etc.[2]

Statistical data obtained can be qualitative (discreet) or quantitative (continuous). Discreet is divided into ordinal and nominal type of data while continuous is divided into ratio or interval scale. Presentation of data can be done through tabulation or drawings. Quantitative data can be presented through graphs that include histogram, frequency polygon, frequency curve, line chart or graph, cumulative frequency diagram, scatter or dot diagram. Qualitative data can be presented through diagrams that include bar diagram, pie/sector pictogram, and map diagram/spot map.[2]

Normal distribution curve is smooth, bell shaped, perfectly symmetrical curve. The total area of the curve is one. Its mean is zero and its standard deviation is one. The mean, median, and mode all coincide. The curve is left negatively skewed when mode is more than median, and right positively skewed when median is more than the mode.[2]

Tools of measurement include rate, ratio, and proportion. Rate measures the occurrence of some particular event in a population during a given period of time, e.g. death rate. Ratio is the relation in size between two random quantities, e.g. male: female ratio. A ratio which indicates the relation in magnitude of a part of the whole is called proportion, e.g. proportional death rate.

Measurements in epidemiology are mortality, morbidity, and disability measurements. Crude death rate, specific death rate, case fatality rate, survival rate, proportional mortality rate, standardised mortality ratio are mortality measures. Morbidity measurements include incidence and prevalence. Incidence is the number of newly diagnosed cases during the period divided by average number of persons at risk during the period multiplied by constant. It is average probability of developing the disease within a time interval, conditional on the absence of disease at the start of the interval. Prevalence indicates all current cases (old and new) existing at a given point of time or over a given period of time in a given population. Prevalence can be point or period prevalence. Period prevalence=point prevalence at the beginning of the interval plus incidence during the interval.

Disability Adjusted Life Years (DALY), Health Adjusted Life Expectancy (HALE), and Sullivan's index are disability measures. DALY is years of life lost to premature death and years lived with disability adjusted for the severity of the disability. Equivalent number of years in full health that a newborn can expect to live based on current rates of ill health and mortality is HALE. Sullivan's index is the expectation of life free of disability, i.e. duration of disability subtracted from the expectation of life at birth. The Indian Disability Evaluation Assessment Scale (IDEAS) for mental illness is a commonly used scale for disability assessment.

Association and Causation

When the disease is multi-factorial, numerous factors or variables become implicated in the web of causation. In other words, events are said to be associated when they occur more frequently together than one would expect by chance. Observed association between a disease and suspected factor not being real is spurious association.

Mean, median, and mode are measures of central tendency. Mean is sum of all observations made divided by number of observations made. Median is the central value of a series of observations. Mode is the most frequently occurring value in a series of values or observations.

Biological data, quantitative or qualitative, collected by measurement or counting are very variable. There are three main types of variability: Biological, real, and experimental variability. Biological variability can be individual, periodical, sampling, and class or group of category variability. Observer error, instrumental error, and sampling defects or errors of bias can result in experimental variability. Range, mean deviation, standard deviation (SD), and coefficient of variation are measures of variability of individual observations. Measures of variability of samples include standard error of mean, standard error of difference between two means, standard error of proportion, standard error of difference between proportions, standard error of correlational coefficient, and standard error of regression coefficient.

Range is the difference between the highest and the lowest figures. Mean deviation is the average of the deviations from the arithmetic mean. SD is the root mean square of the deviation. Coefficient of variation is a measure used to compare relative variability.[2]

Probability is the relative frequency or chance of occurrence with which an event is expected to occur on an average, such as giving birth to a boy in first pregnancy, chances of one drug being better than other, etc. Probability (p)=number of events occurring divided by the total number of trials; q=1-p is the probability of these events not occurring.

Significance is the opposite of chance. The level of significance is expressed as 'p' value (i.e. probability value). Thus p value denotes whether the difference in the sample estimates is due to chance or due to an external factor. Test of significance help us examine the validity of the inference drawn from our observations. The phrase 'test of significance' was coined by statistician Ronald Fisher.[13]

Significance

If the mean of one sample differs from the other sample by more than two times the standard error, i.e. lying outside the 95% confidence limit, the difference is said to be statistically significant at five per cent level (p value less than 0.05; probably due to the play of some external factor and not chance). A value lying outside three standard errors or 99%

confidence limit is very rare and the difference is considered to be highly significant, i.e. p value less than 0.01. Thus, if p<0.05 the test is statistically significant and if <0.01, it is highly significant. Setting the level of significance depends on the study topic. In most statistical studies, the levels of significance are set at five per cent (p<0.05), one per cent (p<0.01), and 0.5% (p<0.005).

Based on the types of the test data, the analysis also varies. Parametric tests assume a normal distribution, and use ratio or interval as typical data and central measure as mean. The non parametric tests assume any distribution and any variance, and central measure is median. Parametric helps draw conclusions while non parametric is simpler, less affected by outliers, e.g. of parametric is independent t-test, while non parametric is Mann Whitney test.

Tests of significance for qualitative data (sample size >30) are standard error of proportion, standard error of difference of proportion, chi-square test; they can also be applied for small samples. For quantitative data, if sample size >30, then standard error of mean and standard error of difference of mean is used. If sample size <30, paired t-test and unpaired t-test is used. When more than two groups are to be tested, analysis of variance (ANOVA) is applied.

Chi-square test involves the calculation of a quantity called chi-square (χ^2). It was developed by Karl Pearson. It is used to compare the observed and expected frequencies, and to test the statistical significance of association between two discrete (qualitative) variables. Categorical data may be displayed in contingency tables. The χ^2 statistic compares the observed count in each table cell to the count which would be expected under the assumption of no association between the row and column classifications. The χ^2 statistic may be used to test the hypothesis of no association between two or more groups, populations, or criteria. Observed counts are compared to expected counts. The χ^2 statistic is calculated under the assumption of no association. Large value of χ^2 statistic $\Rightarrow$ small probability of occurring by chance alone ($p<0.05$) $\Rightarrow$ conclude that association exists between disease and exposure. Small value of χ^2 statistic $\Rightarrow$ large probability of occurring by chance alone ($p>0.05$) $\Rightarrow$ conclude that no association exists between disease and exposure. It also has a role in testing the statistical significance of a linear trend in the magnitude of parametre values in relation to an ordered factor (ordinal variable) and for testing the 'goodness of fit' of a sample distribution of a variable with respect to a standard distribution. Data should be in the form of counts and should be randomly collected. Observations must be statistically independent, sample size >50 (large), linear equation, and frequency of the group more than ten. Yates correction may be applied if observed frequency is less than five. Test does not indicate cause and effect between two relationships. It gives association and not strength of association.

Fisher exact test is used if total sample size is less than 40 and one or more expected frequency is less than five. McNemar's matched chi-square test is used when the data are paired or correlated.[2,7,8]

Z test is normal deviate test, based on normal probability distribution and is used for judging the significance of several statistical measures, especially mean. Assumptions to carry out z test are: Variables are normally distributed, samples are randomly taken, and sample size is sufficiently large (>30).[2,7,8]

The t-test is used when we have a small sample, i.e. sample size <30 and the sample means are not distributed normally about the population mean. In such cases, t-test is used to test the significance. Unpaired t-test is used when the observations are independent. Paired t-test is used when the observations are dependent.[2,7,8]

ANOVA is used when simultaneous comparisons are made of measurements from more than two samples. When measurement data are influenced by several kinds of effects operating simultaneously, this statistical technique is adopted to decide which effects are important, e.g. to test whether means of haemoglobin levels of three groups

Table 1: Parametric and non parametric tests[14]

	Parametric	Non parametric
Assumed distribution	Normal	Any
Assumed variance	Homogeneous	Any
Typical data	Ratio or interval	Ordinal or nominal
Data set relationships	Independent	Any
Usual central measure	Mean	Median
Benefits	Can draw more conclusions	Simplicity, less affected by outliers
Tests		
Choosing	Parametric tests	Non parametric tests
Correlation test	Pearson	Spearman
Independent measures, 2 groups	Independent measure t test	Mann Whitney test
Independent measures, >2 groups	One way, independent measures ANOVA	Kruskal-Wallis test
Repeat measures, 2 conditions	Matched pair t-test	Wilcoxon test
Repeated measures, >2 conditions	One-way, repeated measures ANOVA	Friedman's test

ANOVA=Analysis of variance

of children fed with three different diets are statistically significant.[2,7,8]

Correlation is the method of investigating the relationship between two characteristics, both of which are quantitative in nature. A measure of the strength of relationship between two variables is provided by the coefficient of correlation, denoted by 'r' and termed as Pearson's coefficient of correlation. We define it as the measure of linear association between two quantitative variables. $r=1/n\Sigma(x - \bar{x})(y - \bar{y})/(SD \text{ of } x)(SD \text{ of } y)$. x and y denote the variables, $\bar{x}$ and $\bar{y}$ denote arithmetic means, n is total number of observations. Presence of high correlation signifies linear relationship. Absence does not mean no correlation, it only means no linear relationship exists. r lies between -1 and +1. Positive r indicates positive linear association between x and y or variables, and negative r indicates negative linear relationship. r is always between -1 and +1. The strength increases as r moves away from zero toward either -1 or +1. The extreme values +1 and -1 indicate perfect linear relationship (points lie exactly along a straight line). Graded interpretation: r 0.1-0.3=weak; 0.4-0.7=moderate, and 0.8-1.0=strong correlation. Spearman is used for non-normal distribution, e.g. income and intelligence quotient (IQ). Pearson is used for variables that are normally distributed, e.g. height and weight.

Regression analysis predicts the value of dependant variable (y) when the value of independent variable is given or known (x). Analysis describes the dependence of a variable on an independent variable suggesting possible cause and effect relationship between factors and explains some of the variation of the dependant variable by independent variable by using latter as control. We calculate back, 'regress' the dependant variable based on values of independent variable. While correlation gives the degree and direction of relationship between two variables, regression analysis enables us to predict the value of one variable on the basis of other. Thereby the cause and effect relationship is perfectly understood. Regression coefficient B, also called standardised coefficient, is an estimate that results from regression analysis. It has been standardised, so that dependant and independent variables have a variance of one.

Correlation or regression in medical studies: Correlation measures closeness and strength of linear relationship, when we are not aware of which one of the two causes the other. Correlation is useful to summarise the relationship. But, it cannot quantify the change that can be effected in a variable by changing the other variable that causes the change. If interest lies in quantifying the change in one phenomenon when we change the other, we use regression analysis.

Diagnostic testing: Sensitivity, specificity, and validity are important diagnostic requirements in any study. Sensitivity means the ability of the test to detect true positives, i.e. ability to find out how many people truly suffer from the disease. Specificity is the ability of the test to detect the true negative meaning; how many people in study is truly disease free.

Positive predictivity value answers a vital question that if the person has been tested positive, what is actually the probability that he or she truly has the condition. Similarly, negative predictive value decides if the person has been tested negative, what is the probability that he or she truly does not have the condition? Likelihood ratio of a test is the likelihood of finding a positive or negative test in a person suffering from the condition or not suffering from the condition respectively.[1]

Reliability measures the inherent performance of the instrument or study. The reliable instrument or a study result will produce consistent results when it is applied more than once on the same unit under similar conditions. Test retest, parallel form, inter-observer, internal consistency are types of reliability. Validity is how well a test measures what it is suppose to measure. Types are face validity, criterion-related, formative, and sampling.

Importance of choosing the cut off value: Choosing a higher threshold value decreases both true positive and false positive, increases true negative and false negative. Low value increases true positive and false positive, decreases true negative and false negative. It depends on what we want to choose.

Parts of a Dissertation/Thesis

It includes: Title, abstract, introduction/background, problem statement, purpose/aims/rationale/research questions, review of literature, methodology, significance/implications, overview of chapters, plan of work, and bibliography.

Various statistical software like Epi info, SYSTAT, statistical package for the social sciences (SPSS), statistical analysis system (SAS), STATA, biomedical package (BMDP), and many more are available.

Limitations of statistics include data has to be adequate. Individual items or a few data cannot prove anything. Results are true only on average and deal with average of any population. Statistics cannot reveal the entire story. It deals with numerical statement of facts and cannot analyse qualitative information like honesty, intelligence, beauty, unless transformed into quantitative forms. They are liable to be misused. They are means, not a solution to the problem It is a useful tool; but, very dependent on its interpretation.

Conclusion

While clinical correlation is essential, we cannot negate the importance of statistics in proving our research. One cannot do research now without the basic understanding of statistics. While complete understanding and details into this big topic is beyond the scope of this 'compendium', we have tried to capture the essence of this very modern and vast science.

References

1. Park K. Textbook of preventive and social medicine. 21st ed. Banarasidas Bhanot; 2011.
2. Mahajan BK. Methods in biostatistics. 7th ed. New Delhi: Jaypee Brothers Medical Publishers; 2010.
3. Graunt J. Natural and political observations made upon the bills of mortality. Baltimore: The Johns Hopkins Press; 1939. [This book was originally published in London in 1662]
4. Registrar General of England and Wales. First annual report of the registrar-general of births, deaths, and marriages in England. London: His Majesty's Stationery Office; 1839:99-102.
5. Eyler J. Victorian social medicine: The ideas and methods of

William Farr. Baltimore: Johns Hopkins University Press; 1979.
6. Pelling M. Cholera, fever, and English medicine, 1825–1865. Oxford Press; 1978:92-102.
7. Sundaram KR, Dwivedi SN, Sreenivas V. Medical statistics principles and methods. New Delhi: BI Publications; 2010.
8. Lowe CR, Kostrzewski J. Epidemiology: A guide to teaching methods. London: Churchill Livingstone; 1975.
9. Venkataswamy Reddy M. Statistics for mental health care research. Bangalore: NIMHANS publications; 2002.
10. Cochran WG. Sampling techniques. 3rd ed. New York: Wiley; 1977.
11. Yakoot M. Bridging the gap between alternative medicine and evidence-based medicine. J Pharmacol Pharmacother. 2013;4:83-5.
12. Last JM. A dictionary of epidemiology. 3rd ed. New York, NY: Oxford University Press; 1995.
13. Fisher RA. Statistical methods for research workers. Edinburgh: Oliver and Boyd; 1925.
14. Changing Minds. Parametric vs. non-parametric tests [Internet]. [cited 2015 Oct 30]. Available from: http://changingminds.org/explanations/research/analysis/parametric_non-parametric.htm

Bhuyan D, Dua N, Kothari T. Epidemiology and biostatistics: fundamentals of research methodology. Open J Psychiatry Allied Sci. 2016;7:87-93. doi: 10.5958/2394-2061.2015.00022.1. Epub 2015 Nov 11.

Source of support: Nil. **Declaration of interest:** None.