Annals of Clinical and Medical Case Reports

MiniReview ISSN2639-8109Volume8

Bio-Statistics Newer Advances, Scope & Challenges in Bio-Medical Research

Murthy NS[1], Kunnavil R[2] and Shivaraj NS[3]

[1]Former Research Director, DRP and Professor and Research Coordinator, M.S. Ramaiah Medical College & Hospitals (MSRMC),Bangalore, India.

[2]Lecturerin Biostatistics,National Institute ofUnani Medicine,Bangalore,India

[3]Assistant professor Statistics, Department of Community Medicine, M.S. Ramaiah Medical College & Hospitals (MSRMC), Bangalore, India

***Correspondingauthor:**

Radhika Kunnavil.
LecturerinBiostatistics,NationalInstituteof
Unani Medicine, Bangalore- 560091, India,
Tel: 7829606040,
E-mail:radhik121@gmail.com

Received: 25 Feb 2022
Accepted: 09 Mar 2022
Published:14Mar2022
JShort Name:ACMCR

Citation:

Kunnavil R, Bio-Statistics Newer Advances, Scope & Challenges in Bio-Medical Research. Ann Clin Med Case Rep. 2022; V8(14): 1-7

Keywords:

Biostatistics; Bio-medical research; Scope;Applications; challenges

1. Abstract

Biostatistics also known as biometry which means 'measurement oflife'isabranchofappliedstatisticswhichdealswithcollection, compilation,analysisandinterpretationofdatarelatedtobiomedicalsciences.Itprovidesakeytobetterunderstandingofthemed- ical discipline. Biological data are always subjected to variation andareaffectedbyvariousenvironmental,socialandgeneticfac- tors etc. Biostatistics proves a tool for analysing the data taking into account the variability and to elicit meaningful conclusion in research. In this era of evidence-based medicine. Nowadays, the discipline of biostatistics acts as a basic scientific entity of public health, health services and biomedical research. Biostatistics has evolved as a branch of science in biomedical research that uses a combinationofstatistics,probability,mathematicsandcomputing toresolveperplexingquestionsmathematically.Becauseresearch questions in biology and medicine are diverse, biostatistics has expanded its domain to include any complex quantitative model toanswerresearchquestions.Biostatisticshasreachinwidespec- trum of biomedical domains. From drug development in a labo- ratory to development of various disease control and prevention measuresatcommunitylevelandalotmore.Theimportantaspect forbiostatisticsisthatithelpsinvalidclinicaldecisionmakingby elucidating and quantifying the contribution of chance. Overtime thisfieldhasexpandeditswingsandnowitiswidelyusedacross manysystemsofmedicineincludingAyurveda,Unanietc.The computerprogrammesandadventofnewermethodshasincreased itsscopeandapplications.Thispaperaimstothroughlightonthe newer advances in biostatistics. This paper explores the newer methods, scope & few challenges in Bio-Medical Research.

2. Introduction

Bio-statistics has pivotal role in the development of medical and biological sciences as well as in the development of various disease control and prevention measures. Nowadays, the discipline of biostatistics acts as a basic scientific entity of public health, health services and biomedical research. Over the last few decades, bio-statistics have become more quantitative, stochastic, evidenced based with the growth of medical sciences and public health-oriented research. Emerging disciplines such as Machine learning, Clinical Epidemiology, Molecular Biology, Genomics and Pharmacokinetics have all contributed to making medicaland health sciences depend more and more on Biostatistics. The present write-up focusses on role and use of bio-statistics in the epidemiology,bio-medicalresearchaswellassometotouchupon newermethods,scope&fewchallengesinBio-MedicalResearch.

3. Bio-StatisticsasaStream

Over the last few decades, the development of statistical methods has expanded. Statistics applied to medical research – biostatistics–maynowberegardedasubjectinitsownright,research inmedicineandpublichealthhasbeenbothabenefactoranda source of new difficulties as a result of this new technique. In reality, biostatistics has evolved into a distinct field of study that solves issues in the biological sciences by combining statistics, probability,mathematics,andcomputing.Biostatisticshasbroadened its area to encompass any quantitative, not just statistical, modelthatmaybeusedtoanswertheseissues,duetothediversity of research questions in biology and medicine. Biostatistics is a fieldthataimstoprovideinformation.Consequently,biostatistics draws quantitative methods from fields including statistics, operations-research, economics, and mathematics in general; and it is appliedtoresearchquestionsinfieldssuchasepidemiology,nutrition,environmentalhealth,andhealthservicesresearch,genomics and population genetics, clinical medicine, and ecology.

Thesignificanceofbiostatisticsandbiostatisticiansinmedicalresearch has long been acknowledged by the biomedical community, and statistics in medicine may now be regarded a successful paradigmfortheincorporationofstatisticsintoscientificpractice. Therelevanceofbiostatisticiansinthebiomedicalprofessionmay be seen in the fact that they are frequently asked to contribute as advisersonrenownedcommitteesandjournals.Furthermore,specificstatisticalpublicationssuchasBiostatistics,Biometrics,Biometrika, and many other biostatistics-related journals are held in high esteem [1-2].

4. Types of Research Investigations in Bio-Medical Field

Quality biomedical research is based on a foundation of careful studydesign.Overthelastseveraldecades,newerandinnovative conceptsandstatisticalmethodsforthedesignandanalysisofdata inbiologicalstudieshavebeenestablishedandarebeingused.Designofstudiessuchascase-controlstudies,cohortstudies,clinical trials,andsurvivalstudieshasbeenthecenterofdevelopment.The application of epidemiologic ideas and techniques to the design, conduct, and analysis of clinical trials is a major development, with comprehensive applications described in the following paragraphs. Observational and experimental research studies are the two categories of scientific research studies in biomedical field. Selection of subjects on whom measurements are made is one of the most essential problems that occur during the formulation of statistical methods of research [3-5].

5. WhyisStatisticsNecessaryinBio-MedicalField?

Withoutappropriateinferences,empiricalresearchinanyfield is incomplete, and biomedical research is no exception. Both the designofdiversebiomedicalresearchinvestigationsandtheevaluationofoutcomesneedtheuseofproperstatisticaltools.Biostatistics has now become a crucial component in several research domains as a result of expansion of quantitative approaches with in biomedical sciences (bio-chemical, physiological, clinical parameters, or evidence-based medicine). Medicine is a science in whichchanceplaysanimportantrole.Statisticsasascienceaidin quantifyingtheroleofchance,whereasstatisticsasanartaidsindividualcliniciansinmakingaccuratediagnostic,prognostic,and therapeutic judgments. It also aids health programme administratorsandpolicymakersintheplanning,monitoring,andevaluation ofpublichealthefforts.Ahealthindicatorcanbeusedtodescribe one or more aspects of an individual's or population's wellbeing (quality,quantity,andtime),andalsotodefinepublichealthproblemsataspecificmomentintime,toindicatechangesinlevelsofa population'sorindividual'shealthovertime,todefinedistinctions inpopulationhealth,andtoassesstheextenttowhichaprogram's objectivesarebeingmet.Similarly,validitymetricsincludingsensitivity,specificity,positiveandnegativepredictivevalueareused to evaluate the quality and usefulness of a diagnostic test or to determine the efficiency of a marker in disease diagnosis [3].

EpidemiologyandBiostatistics

Epidemiology is the branch that studies diseases occurrence and its reasons in different groups of people. Epidemiological data is used to design and evaluate disease prevention initiatives, as well astoguidethetreatmentofpatientswhoalreadydevelopeddisease. Biostatisticsandepidemiologyhavehistoricallyhadsuchasignificant link. The early public health experts were doctors basically keentounderstandthepathwhereinailmentseventuateinpopulations,theircauses,aswellastheirinterrelationswithvariousmedicalandnon-medicalaspects.Theseinnovators'challengesincludednotjustthestudyofepidemicsandnon-communicablediseases suchastheconnectionbetweensmokingandlungcancer,butalso the evaluation of therapies. Many had strong analytical reasoning abilitiesandwerewell-versedinstatisticalmethods.Then,begin- ning in the 1930s, epidemiology began focusing upon that study of chronic diseases.The same prospective research strategies that had been so clearly appropriate in the study of infectious diseases became untenable.And it was statisticians, particularly Cornfield and Mantel, who provided a rationale for clarify case-control inference.Withconcernsaboutbiasrelatedtopossibleconfounding factors,biostatisticianshavebecomemoreinterestedingrowingon theprerequisitesforvalidinference.Theyalsobegansearchingat otherareasofepidemiologicalresearch,includingmodelsforevaluatingtheeffectsofpotentialdiseaseriskfactors,suchasdoseresponsemodels.Theseeffectsarequantifiedemployingprobabilisticnotionssuchastheoddsratioorrelativerisk,whichcanbeestimatedappropriatelybasedonthetypeofstudy(case-control,crosssectional,orcohort)usedforeachresearchproj- ect. The large number of statistical methods required in epidemi- ology has led to the publication of numerous books on statisti-cal applications in epidemiological contexts [6-9].

Clinicaltrialsandbiostatistics

Clinicalstudiesareacrucialcomponentofmedicalresearch.Scientific advance can lead to better ways of diagnosing, detecting, and treating diseases and medical conditions as a consequence of theseclinicalstudies.Clinicaltrialsareresearchstudieswhich use human subjects to evaluate novel therapies or drug combinations, modern surgical or radiotherapy approaches, or new proceduresinordertoimproveillnessdiagnosisorqualityoflifeforpatients. Most hospitals now participate in drug testing, which are only started once laboratory investigations show that a new treatment or technique is safe and also hasthe potentialto be more efficient thanexistingoptions.Statisticshavebecomeincreasinglyimport- ant in the field of pharmaceutical development in recent years. Fromplanningthroughconductandinterimanalysistofinalanal- ysis and reporting, statistics is essential at each and every stageof a clinical trial [10-11]. The statistician is typically in responsi- ble for formulating randomization schedules, advising on sample size, establishing framework for deciding treatment differences, and evaluating response rates. In most instances, the statisticians will also act as a liaison with the Independent Data Monitoring Committee. Several novel and recurring challenges in the drug development process require special attention. Ongoing development of statistical methods for handling subgroups in the design and analysis of clinical trials; alternatives to "intention-to-treat" analysis in the presence of noncompliance in randomized clinical trials; methodologies to address the multiplicities resulting froma variety of sources, methods to assure data integrity etc all of which are inherent in the drug development process. These concerns continue to be a source of contention for statisticians workinginthepharmaceuticalindustryacrosstheworld.Furthermore, the engagement of statisticians from all backgrounds continuesto enrich the profession and contribute to social health improvements. Biostatisticians' significant methodological contributions to clinicaltrialsresearchhasledtothedevelopmentofanewjournal, Pharmaceutical Statistics, which was just published in 2002 andisalreadyplacedintheJCRrankingforStatisticsandProbability.

6. Advanced StatisticalAreas of interest in Bio-Medical Field

In addition to routine descriptive and inferential statistics, generalised linear models, survival analysis, and Bayesian methods etc have already had a significant impact on the medical statistics in recent years (in diagnostic, epidemiological and clinical trials contexts). Regression analysis or linear discriminant analysis are statisticalapproachesusedinthebiomedicalprofessiontopredict a dependent variable using additional independent variables/ features or to divide persons into two or more classes of objects or eventsbasedonillnessstatus.Theapproachesoutlinedabovehave aided in the reduction of dimensionality as well as the classifying of people as sick or non-diseased [7-13].

7. Modeling-Approach in Epidemiological Research/ Bio-Medical Filed - Generalized Linear Models

Modeling of health and disease process has been a complex phenomenon. Several models have been employed for the analysis and interpretation of data in the biological field. In the forgoing sections,itisproposedtodescribethreedifferenttypesofgeneral linear models which have been extensively employed as multivariate procedures in bio-medical field viz. (i)Age-period cohort models, (ii) Logistic regression model and (iii) Survival analysis.

Importanceoftime-relatedanalysis

Cancerincidence/mortalityratesfrompopulation-basedregistries (which gather data on all cancer cases in specified areas) give information on regional and temporal variation in cancer risk by personal variables including age, sex, and racial or ethnic groupings."Time,"thethirdelementofanepidemiologicaldescription, provides information on geographic areas and serves as the foundation for determining how effective cancer-prevention methods are.Changesincancerpatternovertimeareofcriticalimportance in cancer control efforts.Age at risk, calendar year, and birth cohortimpactarethemostoftenevaluatedtime-relatedconfounders. The trend analysis helps to understand the question such as how cancer risk has been changing, why and what is likely to happen in future. Cancer trend analysis is important information for the public health and health care planning. Trend analysis reveals informationonthedisease'saetiologyandthesignificantvariancein itsprevalenceacrossdifferentgeographicareas.Cancerincidence/ mortalitytrendscanalsobeusedtoforecastfuturecancerpatterns, whichcanhelpshapefuturepublichealthpolicy.Thesuitabletools forassessingtrendsincancerincidence/mortalitydataincludedata modellingbyage,birthcohort,andcalendartimeperiod."Age-period-cohort (ACP) models" are the name for these models. These models are based on regression models with a Poisson distribution[14].UsingthedatafromtheIndianPopulationBasedCancer Registries for the past twenty-five years, the aforesaid modelling approachwasusedtopredictthechangesintheincidenceofcom- mon malignancies. Some malignancies, such as breast, ovarian, corpus, and uterus, were found to be rising at a rate of around 1-2 percenteveryyear,accordingtotheresearch,andthesimilartrend was seen in women of a younger age range [15-18].

Modellingof the data incase of binary outcome event

When the dependent variable (outcome variable) is binary in nature,suchaswhetheraneventoccursornot,takingvaluesofunity or zero, the assumption required for fitting a multiple linear regressionmodelofthetype $Y=\alpha+\Sigma^k_{i=1}\beta X_i$ isviolatedbecauseit is unreasonable to assume that error distribution is normal. Multiple logistic regression analysis (LR) is used as a multivariate approach to discover the independent predictors of the outcome variable instead of multiple linear regression analysis. The key difference between LR and multiple linear regression models is that instead of utilising the dependent variable as is, we utilise a model based on the dependent variable's logit transformation to meet the required assumptions. As a result, in the LR model, we forecasttheproportionofsubjects(P)whohaveaspecificcharacteristic, or, alternatively, the probability of having characteristics for any combination of explanatory factors [8,19,20]. A dichotomousoutcomevariableislinkedtoasetof"k"knownorsuspected causes(regressionvariables),aswellasprobableconfoundingand effectmodificationvariables,inthismodel.Covariatesorexplan- atory variables are a collection of k regression variables or risk factors. The approach of maximum likelihood estimation is used to estimate the unknown parameters in the model, α, and β_i. In most cases, the likelihood inference is preceded by the fitting ofa hierarchy of models, each of which contains the last variable. ThelikelihoodratiotestoratestbasedonWaldstatisticsareused to test the hypothesis. This modelling approach was used to find independent risk variables linked to illness or a negative outcome event. The logistic regression approach was used in research to lookattheimpactofmaternalandperinataloutcomesindifferent degreesofanaemia.Mildanaemiahadthebestmaternalandperi- natal outcomes, according to the research. Low birth weight kids, induction rates, surgical deliveries, and protracted labour have all been linked to severe anaemia [21].

Studiesonsurvival analysis

Understanding the link between time and the occurrence of vital andhealth-relatedeventsrequiresthestudyoflifetimedata.Inthe biomedical area, time-to-event data is regularly encountered for study.Thistypeofstudyisknownas"survivalanalysis."Thetime passed between a subject's enrollment into the research and the occurrence of an event that is related to treatment is the outcome variable in follow-up/survival studies. The outcome variable has been dubbed survival time, and the event of interest (the onset of a disease) has been dubbed failure. In oncology, for example, the focusisusuallyonthepatient'schanceofsurvivalafterasurgical procedure.Issuesofcensoringandtruncationhampertheanalysis of this sort of survival trial.

Theanalysisoflifetimedataisimportantinunderstandingthere- lationshipbetweentimeandoccurrenceofvitalandhealthrelated events. Time-to-event data is frequently encountered for analysis inbio-medicalfield.Suchanalysisiscalledas"survivalanalysis". In follow-up/ survival studies, the outcome variable is the time elapsed between the entry of a subject into the study and the occurrence of an event is related to treatment. The event of interest (development of a disease, death) has been referred to failure and the outcome variable as the survival time. In oncology, for example, interest typically centers on the patient’s time of survival followingasurgicalintervention.Theanalysisofthistypeofsurvival experiment is complicated by issues of censoring and truncation.

Censoring occurs when we do not fully observe the patient’s survival, due to death unrelated to the cancer under study, or disappearance from the study for some reason. The other factor is truncation, which basically occurs when some patients can’t be observedforsomereasonsrelatedtothesurvivalitself.Acommon example of this is in HIV/AIDS studies of the incubation period (i.e., time from infection to disease). The follow-up starts when the HIV virus is detected and the moment of infection is retrospectivelyascertained.Severalsurvivalparametricmodelssuchas Exponential and Weibull distributions were introduced to model the survival experience/follow-up data analysis of homogeneous populationsincorporatingthecensoringschemes.Thedistribution ofsurvivaltimesmustbeknowntoapplythesemodels.However, whenthedistributionofsurvivalisnotknown,thenon-parametric methodofKaplan-Meiercurvedevelopedin1959hasbeenawell- knownestimatorofthesurvivalfunction,anditisextensivelyused in epidemiological and clinical research [20-24].

In order to take into account diversity of situations, which were encountered in practice, Cox in 1972 developed a modelling procedure termed as Cox-proportional hazards model under a very rigorous theoretical backup. The classical proportional hazards model of Cox (1972) is also widely used whenever the goal is to study how covariates affect survival. This model is an important tool in the follow-up/survival studies for modelling the effect of riskfactors/prognosticfactorswhentheoutcomeofinterestoccurs with time. In the model, the hazard for an individual is a part of the product of a common baseline hazard and a function of set of risk factors. By applying the above modelling procedure, the independentriskfactorsassociatedwiththedevelopmentofprecan- cerous lesions of cancer of cervix was evaluated [25]. Similarly, in another study the treatment effectiveness for curing of a gastro intestinal bleeding was evaluated which employed an experimentaldesign [26].However, when the assumption of proportionality does not satisfy, then a classical approach for the analysis of data ofthistypeisthetime-dependentCoxregressionmodel(TDCM). Advantages of Cox’s regression model include its easy interpretability and its availability in the majority of statistical packages.

Newissuesinsurvival analysis

When survival is the ultimate result yet intermediate phases are discovered, a generalisation of the survival process occurs. Inthis case, a series of occurrences is witnessed, resulting in many observations per person. Intermediate stages might be based on categorical time-dependent factors like transplantation, clinical symptoms (e.g., bleeding episodes), or a complication during the disease (e.g., metastases), or biological markers (like CD4T-lymphocytelevels).Multi-statemodels(MSMs)wereaccessibleinthe 1990s,allowingforabettergraspofthediseaseprocessandabet- ter comprehension of how the time dependent covariate impacts theillness'sevolution.ComparedtoCox'sregressionmodel,these contemporary models offer significant advantages. They provideabetterunderstandingoftheillnessprocessbyindicatingtheris k ofmovingfromoneconditiontoanother(transitionintensities),as wellasavarietyofadditionaldata,suchastheaveragetimespent in each state and survival rates for each stage. Differences in the course of sickness among subjects in the population can also be explainedbycovariatesontransitionintensities.MSMs,inpartic- ular, can show how various variables effect different transitions, which is impossible to do with other models like the TDCM. In reality,theriskofmortalityinindividualswhohaveundergone differenttherapyisunlikelytobethesame.Furthermore,prognos- tic markers linked to the risk of mortality may vary based on the treatment received, for example. There is currently a substantial amount of research accessible on the analysis of MSMs [27-31].

8. SomeRecurrentandEmergingIssuesin Biostatistics

In terms of both the continued improvement of traditional approaches and the introduction of new techniques to meet new issues, modern biostatistics faces a variety of obstacles. We next turnourfocustoanumberofemergenttopicsthatbiostatisticians shouldinvestigatefurther,includingbioinformatics,spatialstatistics, neural networks, and functional data analysis, as well as big data analysis.

Statisticalmethodsinbioinformatics

Averyrapidlyemerginginfluenceonbiostatisticsistheon-going revolutioninmolecularbiology.Molecularbiologyisnowevolv- ing towards information science, and is energizing as a dynamic newdisciplineofcomputationalbiology,sometimesreferredtoas bioinformatics. Bio-informatics merges recent advances in molecularbiologyandgeneticswithadvancedstatisticsandcomputer science. The goal is increased understanding of the complex web of interactions linking the individual components of a living cell totheintegratedbehaviouroftheentireorganism.Theavailability of large molecular databases and the decoding of the human genomemayallowascientisttoplananexperimentandimmediately obtain the relevant data from the available databases. This is an area in which statistical scientists can make very important contributions. Several biostatistics departments (mainly in the U.S.) have already been renamed as “Biostatistics and Bioinformatics” [32-33].

Spatialstatisticalmethodsinhealthstudies

In numerous types of public health and epidemiological studies, the investigation of the geographical distribution of illness incidence and its link to possible risk factors plays an essential role. Geographic epidemiology is the overall term for this field, and there are four major statistical areas of interest: (a) Given "noisy" observeddataonillnessrates,diseasemappingtriestoconstructa mapofthegenuineunderlyinggeographicaldistributionofdisease incidence.

(b)Ecologicalstudieslookforcorrelationsbetweensicknessincidence and potential risk factors in groups rather than individuals, withgroupsfrequentlydefinedbygeographiclocation.Suchstud- ies are helpful in discovering the cause of sickness and may contribute in suggesting future research paths as well as prospective preventativestrategies.(c)Diseaseclusteringresearchfocuseson finding geographical locations with a considerably higher risk of disease, or evaluating the evidence of heightened risk near potential sources of hazard. The exploration of control measures when the aetiology of observed clustering has been established, or the targetingoffollow-upstudiestodetermineexplanationsforobserved clustering in disease incidence. (d) Environmental assessmentandmonitoringisconcernedwithdeterminingthegeograph- ical distribution of health-related environmental elements and exposure to them in order to develop appropriate controls or take preventative action. Given the scope and relevance of the issues raised by spatial epidemiology, it's no surprise that there's been a lot of interest in this field in recent years [34-36].

Neuralnetworksinmedicine

Manyresearchershavebeendrawntoneuralnetworks(NN)techniques in medicine, and these approaches have been used in a variety of biomedical applications, including diagnostic systems, biochemical analysis, image analysis, and drug discovery. Neural networks,whichmimicthebehaviourofhumanneuronnetworks, havethepotentialtobebeneficialinawiderangeofapplications. NNs, unlike humans, are not influenced by factors like as weariness,workingenvironment,oremotionalstate.NNsarefrequently employed in diagnostic systems, for the diagnosis of cancer and heart issues, and for the analysis of many types of medical pictures (such as tumour detection in ultrasonograms, classification of chest x-rays, and tissue and vascular classification in magnetic resonance imaging). Many researchers are interested in neural networks (NN) approaches in medicine, and these NNs are being used experimentally to model the human cardiovascular system: diagnosis can be achieved by building a model of an individual's cardiovascularsystemandcomparingittoreal-timephysiological measurements taken from that patient. NNs are also employed in the research and development of cancer and AIDS medications.In addition to classical and current statistical approaches, neural networksareincreasinglybeingviewedasanextensiontogeneric statistical methodology [37-39].

Functionaldataanalysisandmedicine

Because of technology advancements in recent years, many scientific domains including applied statistics are increasingly measuringandrecordingcontinuous(i.e.,functional)data.Manycur- rent devices, for example, allow biomedical researchers to obtain functionaldatasamples(mainlyascurves,thoughalsoasimages). Becausefunctionaldataisdisplayedasacurve,thecurveisagood startingpointforfunctionaldataanalysis.Functionaldataofteninvolves a large number of repeated measurements per subject, and these measurements are typically captured at the same (generally similarly spaced) time intervals and with the same high sampling rate for all participants. The derivatives of these curves, as wellas the positions and values of extremes, are occasionally of interest. In endocrinology, for example, investigations of hormone levels after various pharmacological dosages; or in neuroscience, for example, studies to estimate the firing rate of a population of neurons,wheretheunitofresearchiseachindividualneuron'sfiringcurve.Anotherexampleisthestudyofgrowthcurvesinwhich manycharacteristicsofgrowth,suchasheightandlungfunction, arestudied[40].Thegoalsoffunctionaldataanalysisaregenerally exploratory in nature, with the goal of representing and displaying data in order to highlight noteworthy qualities that may then be used as input for further research. Other goals might include estimating individual curves from noisy data, characterisation of homogeneityandpatternsofvariabilityacrosscurves,andevaluationofthecorrelationsbetweencurveformsandvariables.Despite substantialrecentadvancesinfunctionaldataanalysis,thestatisticalcommunityhasahugeproblemindevelopingnewtoolstodeal withfunctionaldata.Asidefromthebooksalreadymentioned,the specialissueonfunctionaldata,whichissettoappearinthejournal Computational Statistics and DataAnalysis soon (2007), may beanexcellentplacetostartthinkingaboutnewareasofstatistical study and prospective applications.

9. New Statistical Methods which are Likely to Play a Key Role in Biomedical Research Over Coming Years

Thefollowingnewstatisticalmethodsarelikelytoplayakeyrole in biomedical research over coming years: (i) bootstrap (another computer-intensive methods); (ii) Bayesian methods); (iii) generalized additive models;(iv) classification and regression trees (CART);(v)modelsforlongitudinaldata(generalestimatingequations);and(vi)modelsforhierarchicaldata,(vii)bigdataanalysis [41]. Modern health research involves increasingly sophisticated statistical tools and computerized systems for data management and analysis. During the past few years' tremendous amount of softwarehasbeenmadeavailabletosupportstatisticalcomputing requirementsforbiomedicalresearch.Bio-statisticianshavetobe extremelyfamiliarwithvariousstatisticalsoftwarepackagessuch as STATA, SAS, SPSS,R etc.

The traditional component of biomedical courses will probably focus on areas of mathematical statistics including probability theory, inference, re-sampling methods (e.g. bootstrap), linear regression,analysisofvariance,generalizedlinearmodels,survival analysis (including multi-state models), nonparametric methods, anddataanalysis.Inaddition,newmethodologieslikespatialstatistics, neural networks, smoothing regression methods (such as generalizedadditivemodels)andoperationsresearcharestrongly recommended. The decision technologies, tools and theories of operationsresearchandmanagementScienceshavelongbeenapplied to a wide range of issues and problems within health care.

10. Conclusion

Biostatistics is a fundamental scientific field in public health, health services, and biomedical research. With the rise of medical sciences and public health-oriented research over the last few decades,biostatisticshasgrownmorequantitative,stochastic,and evidence-based.Emergingfieldsincludingmachinelearning,clinical epidemiology, molecular biology, genomics, and pharmacokineticshaveallledtoagrowingrelianceonbiostatisticsinmedicalandhealthsciences.Medicineisascienceinwhichchance is a significant factor. Statistics as a science may help quantifytheinfluenceofchance,butstatisticsasanartcanhelpindividual doctors make appropriate diagnostic, prognostic, and therapeutic decisions. The use of Biostatistical methods to address issues in clinical trials, survival analysis, Data modelling using Generalizedlinearmodels,genetics,ecologyandmachinelearningetcare gaining much popularity in the present era of in epidemiological biomedical research.

References

1. AltmanDGandGoodmanS.Transferoftechnologyfromstatisticaljournals to the biomedical literature: Past trends and future predictions.JournalofAmericanMedicalAssociation.1994;272(2):129-132.
2. SuarezCC,ManteigaWG,CienciaA.PensamientoyCultura CLXXXIII 725 mayo-junio. 2007; 353-361.
3. Hennekens CH, Buring IE. Design strategies in epidemiologic research.In:MayrentSL(ed).Epidemiologyinmedicine.Boston:Little, Brown and Co. 1987; 16-20.
4. MathewA,Murthy NS.Astep towards quality medical research. National Medical J India. 1998; 11(6): 283-286.
5. Prentice RL. Statistical methods and challenges in Epidemiology and Biomedical Research, Handbook of statistics, Elsevier, 2008.
6. Rao NSN, Murthy NS. Applied statistical methods in Health Sciences, JAYPEE Brothers Medical Publishers (P) LTD, New-Delhi, 2008; 2nd ED, Pg no-1-246.
7. ArmitageP,Statisticalmethodsinmedicalresearch.Oxford:BlackwellScientificPublications.1983;167-188.
8. Breslow NE, Day NE. Statistical methods in cancer research. The analysis of case-control studies. InternationalAgency for Research on Cancer, Lyon, 1980.
9. Breslow NE, Day NE. Statistical methods in cancer research. The analysis of cohort studies. International Agency for Research on Cancer, Lyon, 1987.
10. Murthy NS, Mathew A, Sharma JB. Design of research studies in clinicalObstetricsandGynaecology:Obtetrics&GynaecologyToday. 2001; 7(1): 379-386.
11. PetoR,PikeMC,ArmitageP,BreslowNE,CoxDR,HowardSV,etal. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Part II:Analysis and Examples.British Journal of Cancer. 1977; 35(1): 1-39.
12. Armitage P, Berry G. Statistical Methods in Medical Research (3rd edition). Blackwell. 1994.
13. Houwelingen (van) HC. The future of Biostatistics: Expecting theunexpected". Stat Med. 1997; 16(24): 2773-2784.
14. ClaytonDG,Schifflers,E.Modelsfortemporalvariationsincancer rates. II:Age-Perid-cohort models.
15. MurthyNS,BurraU,ChaudhryK,SaxenaS.Changingtrendsinincidence of breast cancer-Indian Scenario. Indian Journal of Cancer.2009; 46(1): 73-74.

16. MurthyNS,ShaliniS,SumanG,PruthvishP,MathewA.Changingtrend sinincidenceofOvarianCancer-IndianScenario.AsianPacJCancer Prev. 2009; 10(6): 1025-1030.

17. MurthyShaliniP,NandakumarBS,SumanG,PruthvishS,Mathew A. Estimation of trends in incidence of cancer of corpus uteri - Indian Scenario. European Journal of Cancer Prevention. 2010; 20,1: 25-32.

18. Shalini CN, Murthy NS, Shalini S, Pruthvish S, MathewA.TrendsinRectalCancerIncidence-IndianScenario.AsianPacificJCancerPrev. 2011; 12(8): 2077-2082.

19. SchlesselmanJJ.Case-ControlStudies.NewYork:OxfordUniver- sity Press 1982.

20. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations.JournaloftheAmericanStatisticalAssociation1958; 53: 457-481.

21. MalhotraM,SharmaJB,BatraS,SharmaS,MurthyNS,AroraR,etal.Maternallandperinataloutcomeinvaryingdegreesofanaemia.Inter J Gynae & Obstet. 2002; 79(2): 93-100.

22. Cox DR. Regression models and life tables. Journal of the Royal Statistical Society. 1972; B34: 187-220.

23. KalbfleischJD,PrenticeRL.StatisticalAnalysisofFailureTime Data.NewYork:Wiley,1980.

24. MathewA, Pandey M, Murthy NS. Survival analysis, caveatesand pitfalls. Eur J Surgical Oncology. 1999; 25: 321-329.

25. MuthryNS,MathewA,YoeleBB,HappanenN.Cohortdataanalysis:IslogisticregressionorCoxproportionalhazardmodelorPois- son regression model? Ever green problems in Epidemiology (Ed) LeenaTenkanen,Publication3,SchoolofPublicHealth,University of Tampere, 1999.

26. Murthy NS, SehgalA, Satyanarayana L, Das DK, Singh V., GuptaMM, et al. Risk factors related to biological behaviour of precan-cerouslesionsofuterinecervix.BrJCancer.1990;61(5):732-766.

27. Sarin SK, Lamba GS, Kumar M, Mishra A, Murthy NS. A comparisonofendoscopicligationandpropronololfortheprimarypre-vetioin of varical bleeding. New Engl J Medicine. 1999; 340(13):988-993.

28. Klien JP, Moeshberger ML. Survival analysis techniquesfor censored Truncated data,Springer Science Business India, 2006.

29. MoeshbergerML,KlienJP.Statisticalmethodsfordependentcompeting risks, Lifetime Data Anal. 1995; 1(2): 195-204.

30. Klien JP, Moeshberger ML. Independent or dependent competing risks: Does it make difference. Tyler and Francis. 1987.

31. Andersen PK, Borgan O, Gill RD, Keiding N. Statistical Models Based on Counting Processes, Springer, New York, 1993.

32. HougaardP.AnalysisofMultivariateSurvivalData,Springer,New York, 2000.

33. VapnikV.ThenatureofStatisticallearningTheory,Springer-verlag, 1995.

34. Ewens WJ, Grant G. Statistical Methods in Bioinformatics (Statistics for Biology and Health), Springer, New York, 2005.

35. CressieN.StatisticsforSpatialData.JohnWiley&SonsNewYork, 1993.

36. SpatialPredictionandKriging,WileySeriesinprobabilityandStatistics for Spatial Data Wiley Serie4s, in Probability and Statistics, 1993.

37. Ripley BD. Pattern Reconnition and Neural Network, Cambridge University Press, 1996.

38. RadhikaKl,GeorgePS,MathewBS,MathewA.Comparisonofartificialneuralnetworkwithlogisticregressionasclassificationmodels for prediction of breast cancer patients'outcome. International conference in Epidemiology; 2012.

39. RipleyBD.Statisticalaspectsofneuralnetworks.InBarndorf-NielsonOE,JensenJL,KendallWS(Eds).NetworksandChaos-StatisticalandProbabiliticaspects,London,Chaman&HillPress,1993.

40. RamsayJ,SilvermanB.AppliedFunctionaldataAnalysis,2ndEdition, Springer, New York.

41. KunnavilR,MurthyNS.HealthcareDataUtilizationfortheBettermentofMankind-AnOverviewofBigDataConceptinHealthcare. InternationalJournalofHealthcareEducation&MedicalInformatics. 2018; 5(2): 14-17.