Invited Talks:

December 1, 11.35 am -12.20 pm

Regional IBD Analysis (RIA): a new method for linkage analysis in extended pedigrees using genome-wide SNP data

Heather J Cordell, Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle upon Tyne NE1 3BZ, United Kingdom.

Abstract: The study of rare variants has revived interest in linkage analysis. However, exact calculations for traditional linkage analysis are computationally impractical in large, extended pedigrees. Although simulation-based methods can be used, they require significant computation and are not exact. We propose Regional IBD Analysis (RIA), a non-parametric linkage method based on comparison of locally and globally estimated identity by descent (IBD) sharing in affected relative pairs (ARPs). In RIA, genome-wide SNP data are used to calculate the “global” expected IBD sharing probabilities for each ARP, against which a “local” set of IBD probabilities, estimated using SNP data within windows of pre-specified width, are compared. The global and local IBD probabilities are used to construct a non-parametric maximum likelihood statistic (MLS)-like test of linkage in each window. We illustrate our method with real nuclear-family data from a study of vesicoureteric reflux and simulated data based on large extended pedigrees from a study of visceral leishmaniasis. RIA successfully detected the linkage signals with significant reduction in computational time (e.g. 2 hours vs 66 hours on 3,626 individuals from 308 extended families, genotyped at 545,433 SNPs) compared to traditional methods. RIA should be useful in studies involving large extended families, in which traditional linkage analysis is not feasible. Additionally, because it does not rely on prior knowledge about familial relatedness, RIA has an additional advantage of being robust to pedigree misspecification and can be used even in the absence of pedigree information.

Keywords: regional IBD analysis (RIA), non-parametric linkage analysis, identity-by-descent (IBD), affected-relative-pair method, extended family

December 2, 9.45 am -10.15 am

MicroRNA Target Prediction

Sanghamitra Bandyopadhyay, Indian Statistical Institute, Kolkata.
Abstract: MicroRNAs (miRNA) are small RNA molecules that perform post translational modification of target messenger RNAs. Prediction of miRNA targets using machine learning approaches is an important area of research. The prediction problem has often been formulated in a supervised classification framework, using known target pairs and non-target pairs for the purpose of training. However, most of these methods suffer from either high false positive or false negative rates. One reason for this is the marked deficiency of negative examples or miRNA-mRNA non-target pairs. In order to overcome this limitation, our group has developed Target Miner, a microRNA target prediction algorithm with systematic identification of tissue-specific negative examples.

In this talk we will first explain briefly the biogenesis of miRNAs. An existing approach of classification based miRNA target prediction method will be described. Thereafter, we will explain our approach of systematically identifying a set of biologically relevant, tissue specific non-target miRNA-mRNA pairs. This involves expression profiling of both miRNAs and mRNAs, miRNA-mRNA structural interactions and seed-site conservation. These high-throughput tissue-specific negative examples and a set of experimentally verified positive examples are then used to build Target Miner, a support vector machine (SVM)-based classifier. Comparative cross-validation results as well as performance on a completely independent experimental test dataset demonstrate the superiority of Target Miner vis-a-vis ten existing target prediction algorithms. Results will also be provided to demonstrate that the superior performance of Target Miner is indeed due to the proposed approach of negative data selection. Finally the gain in performance of an existing method when trained with the proposed negative data will be demonstrated.

December 3, 9.45 am-10.15 am

Analyzing cancer mortality in small areas using Bayesian B-splines

Maria Ugarte, Department of Statistics and OR, Public University of Navarre.

Abstract: In recent years, models incorporating splines have been considered for smoothing risks in disease mapping. Although these models are very flexible, they can be computationally demanding in certain cases. In this talk, several models including one, two, and three dimensional B-splines (penalized or un-penalized) are proposed to analyze spatio-temporal risks. As computing time could be a limitation in real practice, integrated nested Laplace approximations are considered for model fitting and inference. The complete set of proposed models is used to analyze cancer mortality data in Spain.
Regular Talks:

December 1

Session (2.20-3.20)

1. Survival Analysis and Its Application

**Paper Name:** Exploring the Role of Ca 15-3 Markers in Breast Cancer Patients Using Time Dependent Cox Ph Model

**Author:** Prafulla Swain

**Abstract:** The primary objective of this paper is to study the significance of CA 15-3 as a disease marker in monitoring and evaluating the diseases progression of breast cancer patients using a time dependent Cox Proportional Hazard Model. The overall survival was found to be 57.3%. The predictors viz., age, tumor size, tumor grade, lymph nodes and CA 15-3 (time dependent) are found to be significantly associated with the hazard of death. Patients with tumor grade II has significantly higher hazard as compared to patients with tumor grade I (HR= 6.73, 95% CI; 1.91, 47.80), and the CA 15-3 markers (≥25 u/ml) category patients had an elevated hazard of death as compared to patients with CA 15-3 marker (< 25 u/ml), (HR= 2.16, 95% CI; 1.02, 17.65).

**Paper Name:** Fitting Parametric Survival Models With Covariates to Complex Survey Data with Application to Infant Mortality in India

**Author:** Palash Malo

**Abstract:**
Background: Survival Analysis has the ability to maximise the use of available data by allowing the censored observations to contribute to the analysis of time. Most of the researches were focused mainly on either directly to the modelling and analysis of non-parametric or semi-parametric method than parametric method. Accelerated Failure Time (AFT) model is a parametric method to address the same is considered for analysing the contribution of the selected demographic factors for infant mortality using national survey data in Indian context.
Objective: To identify the best parametric survival model with covariates for the national survey data in order to describe the relationship between the covariates and the survival time of infants.

Materials & Methods: We used the data obtained from third National Family Health Survey (NFHS-3) conducted in India during 2005-06. Birth history data was collected retrospectively for five years preceding the survey. A total of 51,555 live births and 2,508 deaths of infants between birth and 11 months of age were recorded during this period and were included in this study. Variables such as mother’s age at the time of birth; mother’s schooling; sex of child; birth order and previous birth interval; and place of residence were also considered as covariates in this analysis. We also assessed time ratio for each of the covariates to comment whether a covariate accelerates or decelerates the time to death.

Analysis of survival time was carried out for Exponential AFT, Weibull AFT, Log-normal AFT, Log-logistic AFT and the Extended Generalized Gamma AFT models. The Newton-Raphson iterative procedure was applied to obtain the MLEs of each of the AFT models. Further the selection of the best parametric survival model was done using AIC and BIC indices.

Paper Name: Bayesian Accelerated Failure Time and Its Application in Chemotherapy Drug Treatment Trial

Author: Atanu Bhattacharjee

Abstract: The Cox proportional hazards model (CPH) is normally applied in clinical trial data analysis, but it can generate severe problems with breaking the proportion hazard assumption. An accelerated failure time (AFT) is considered as an alternative to the proportional hazard model. The model can be used through consideration of different covariates of interest and random effects in each section. The model is simple to fit by using OpenBugs software and is revealed to be a good fit to the Chemotherapy data.

Paper Name: A method of obtaining inter-arrival distribution of consecutive cardiac attacks.

Author: Kalpana Singh, Suddhendu Biswas, Dorairaj Prabhakaran

Abstract: Objective: To explore the inter arrival time distribution between consecutive heart
attacks by using Geiger Muller counter model type I with random variable dead time and palm probability distribution.

Introduction:
Events such as sickness, heart attacks, failure of a machine, pregnancies, etc. are followed by the duration of after effect or ‘happenings’ (Takacs, 1951); that is a gestation or rest period for the force that causes the events. This is known as ‘dead time’ which may be either fixed or a random variable following some probability distributions.

Methodology:
The present exercise is devoted to finding out the distribution of intervals between two consecutive heart attacks. As such events vary between individuals therefore we propose a compound distribution model for the same. Again because of compounding the distributions renewal structure of the process is destroyed. Therefore, for obtaining the inter arrival time distribution between consecutive heart attacks the palm probability (conditional probability of number of events in an interval given that at the beginning of the interval t=0 an event has occurred) have been used. This is used for in predicting the number of events (like births, accidental shocks during a fixed period of time and waiting time distribution between any two events say iᵗʰ and jᵗʰ.

Results:
By using counter model theory the expected number of heart attacks during fixed periods of time have been illustrated. The number of heart attacks under the weighted Poisson process with the intensity of dead time distribution $\mu$ and $\lambda$ for the events are given by

\[ E(N_1)=0.08, \quad E(N_2)=0.25 \quad \text{and} \quad E(N_3)=0.45 \]

\[ \text{var}(N_1)=0.09, \quad \text{var}(N_2)=0.23 \quad \text{and} \quad \text{var}(N_3)=0.37 \]

Inter-arrival distribution between two consecutive heart attacks have been obtained by using palm probability.

Conclusion:
It has been verified that the distribution of time between the first and nth arrival given that first attack occurred at t=0 by using counter model theory is the same as inter arrival distribution between the first and nth arrival of shocks by palm probability.

Paper Name: Rough Set Based Rule Induction Approach for Survival Analysis.

Author: Perumal Venkatesan

Abstract: The attribute-value boundaries in medical diagnosis and treatment are usually vague and imprecise. Physicians diagnose a patient and decide what is the best way to cure them with vague information on signs and symptoms. Rough set is a leading soft computing method and its theory provides methods for knowledge
In this paper, a rough set based new approach for survival analysis is proposed. The decision problem is formulated over database on spinal tuberculosis patients treated under a randomized trial. The rule based method expresses the difference between the expected outcomes under the different treatment groups. The rules for different survival tendencies are also formulated. The new framework is combined with semi-parametric survival models to identify survival patterns. The efficiency of the method is demonstrated using clinical trial outcomes.

2. Demography

**Paper Name:** Impact of Pre Natal Diagnostic Technique Act implementation on child sex ratio in India:

**Author:** Arumugam ISAI

**Abstract:**

**Background:**
The child sex ratio (CSR) in India was chronologically and progressively declining in India since 1961 from 976 to 914 /1000 males in 2011. The liberalized MTP Act 1971 came into existence from 1st April 1972. The sex selective MTPs are one of the major cause for declining trend for CSR. To prevent this, the GOI introduced Pre Natal Diagnostic Technique (PNDT) Act 1994 operationalized from 1st January 1996. Subsequently by rectifying the objections, the Act came into effect from 14th February 2003.

**Aim and Objectives:**
1. To determine the major state wise net decrease of CSR from 1991 to 2001 and from 2001 to 2011.
2. To determine the relationship of foeticide during 2005 to 2010 on the 2011 CSR.

**Methodology:**
The CSR net decrease of major states for the census years 1991, 2001, and 2011 was calculated and average decadal net decrease compared for significance. The total reported foeticide crimes during 2005 to 2010 were collected and correlated with 2011 CSR.

**Results:**
The mean net reduction of CSR from 1991 -2001 was 23.2 ± 22.4. And from 2001 to 2011 was 6.3±17.6. The reduction between the two decades was significant (P<0.05). The reported foeticide crimes were negatively correlated with 2011 CSR.
Discussion:
The association between foeticide and CSR shows negative correlation. The significant reduction of CSR is attributed to the impact of implementation of PNDT Act 1994.

Conclusion:
The study reveals that the illegal practice of sex selective abortions has been curtailed by enforcement of PNDT Act 1994 as foeticide is negatively correlated with CSR.

Key words: PNDT Act- Foeticide - impact- child sex ratio.

Paper Name: A Model Base Approach to Study the Age at Marriage and First Birth in Case of Caesarean Delivery in Uttar Pradesh, India

Author: Brijesh Singh, Shweta Dixit

Abstract: The caesarean delivery is a common practice in present scenario in most of the developed as well as many developing countries including India. Caesarean birth has increasing noticeably over the last years especially in countries with high or middle income. The present study is based on analysis of data from National Family Health Survey third round in India, which investigate the role of age at marriage and age at first delivery on caesarean cases of first birth. Finding suggests that proposed distribution is appropriate for the age at marriage and first birth of the women, whose first birth is caesarean. Maximum likelihood estimates obtained for the parameters of the distribution. The model suggests the modal age at marriage and first birth in case of first caesarean birth in Uttar Pradesh is 18.41 and 20.56 years.

Paper Name: Inequality in the Likelihood of a Male Birth to the Women for Fixed Parity: A Probability Model Approach

Author: Brijesh Singh, Sonam Maheshwari

Abstract: Uttar Pradesh is a hub state for son preference in India due to presence of strong patriarchal belief that influences fertility decisions. In this study an attempt has been made through the modelling of the pattern of male children in Uttar Pradesh, where family size and sex composition are dominated by strong son reference. Initially the model is proposed under the assumptions that probability of male birth remains constant across the population of women and also across the successive births for the same women. Further the model is modified assuming the
concept of clustering. This study enables us to give idea about the risk in contexts of strong son preferences in Uttar Pradesh. In case of lower parity women the chance of son preference is more than in case of higher parity women.

**Paper Name:** A study of morbidity pattern among elderly population in urban India

**Author:** Akhilesh Yadav, Prof. Chander Shekhar

**Abstract:**

Background: The urban area is at risk of a different type of morbidity like chronic lung disease, depression, angina, diabetes, etc. This study has taken up with the objective of exploring the age pattern of morbidity, level of substance use among elderly, and relationship between substance uses among the elderly population in urban India to improve our health care services. Data and Methods: The data on morbidity has been obtained from the SAGE, 2007, which covered 10600 households across the six states. For analyzing of data, bivariate and multivariate have been used by SPSS software.

Results: Morbidity like hypertension, depression, diabetes and disease-related to oral health is highly prevalent in the age group 50-59. However, cataracts, asthma, and stroke are more frequent in the age group of 60-69, while chronic lung disease is more prevalent in the age group of 70-79. Elsewhere, in urban India, around every second people is a smoker in the age group 50-59. The higher percentage of substance use (alcohol) is in the age group 60-69, while tobacco consumers and alcohol drink are elevated in the age group 70-79. Here, Assam is the state where the prevalence of smoking is highest whereas alcohol consumption is highest in the Rajasthan and lowest in Assam. The overall, tobacco consumer is highest (35 %) in urban India followed by tobacco and alcohol consumer (nearly 10 %). Conclusions: In urban India, the highest tobacco consuming state is Assam while among the elderly population; the third most powerful killer diabetes disease is highest in the age group of 50-59 years. Age, tobacco consumption and improvements in wealth quintile are significant risk factors associated with the presence of morbidities like diabetes, hypertension, and stroke. Asthma and oral health are significant increases along with age. Stroke is significantly high among the people of 60-80 years. It is significantly high in the middle and richest households.

**Paper Name:** Probability Projection of Total Fertility Rate for Assessing the Replacement Level of Fertility of Selected State Using Mcmc Technique under Bayesian Modelling

**Author:** Abhishek Chauhan
Abstract: Fertility is the key indicator that gives us an idea about the size and composition of the population; it is one of the most important components of population growth and population dynamics. The Total fertility rate is a key component in population projection; it is the average number of children which a women of hypothetical cohort would bear during her life time if she were to bear children throughout her life at the rate specified by the age specific fertility rate for the particular year if none of them dies before crossing the age of reproduction. This paper is an attempt to use of Bayesian methodology for projection of total fertility rate (TFR). It provides projection of selected state of India namely Uttar Pradesh, Bihar, Assam, Rajasthan, Haryana, Gujarat Madhya Pradesh. The projection has been done by using Gompertz model. Parameters of the model are estimated using MCMC (Monte Carlo Markov Chain) technique in Bayesian procedure. The data is taken from Sample Registration System (SRS) of India. It includes the data for TFR of each considered state since 1981 to 2013. We have assumed non-informative prior distribution for estimating parameter in Bayesian approach. The entire analysis was done by using WinBUGS (Bayesian Inference using Gibbs Sampling for Windows) software.

3. Epidemiology -1

Paper Name: Study on nutritional status and utilization of Anganwadi services among the selected Anganwadi Centres of Patancheruvumadal of Medak district.

Author: Gangadhar Kalapala

Abstract:

AIM:
The Study aim is to assess the health profile and nutritional status of anganwadi children below six years of age.

OBJECTIVES:
1. To assess the nutritional status of anganwadi children aged one to five years of age in the Patancheruvu madal of Medak district.
2. 1.To assess the nutritional status of children below 6 years of age who are attending anganwadi regularly ,irregularly, and not attending anganwadi centres at selected anganwadi centres .
3. 2. To identify and notify the prevalence of malnutrition in children below 6 years of age at selected anganwadi centres of Patancheruvu.
METHODOLOGY:
A cross sectional, community based study was carried out among the 100 anganwadi children in selected anganwadi centres in Patancheruvu and surrounding areas between 0-6 years of age by taking both males and female children equally. Study design: The present study is community based observational, cross sectional study. Study setting: The study was conducted in ICDS anganwadi centres which are located in Patancheruvu.

RESULTS AND DISCUSSION:
By taking the sample of hundred children from the selected anganwadi centres, in the present study the findings have shown that Mean age of the children= 43.27 months Range=36 months (min=24 max=60) Males 50% and Females 50% , Average age of weaning=9.89=10 months, per cent of children attending anganwadi centre =97.6% , Children attending anganwadi centres regularly=53% Children attending anganwadi centre irregular=29% Children attending anganwadi less than two days in a week=18%, these are the major findings related to the percentage of children attending, not attending the anganwadi centre . for the estimation of malnutrition status the z scores or standard deviations were used. In the present study, findings have shown that the prevalence of underweight, stunting and wasting was 35%, 29%, and 7.7% respectively among anganwadi children of Patancheruvu madal of Medak district.

---

Paper Name: Identification risk factors of hypertension in young adults (25-45yrs) of Pune city and develop mini questionnaire as a screening tool.

Author: Pooja Gaigaware, Amol Mankar

Abstract:
Background-Non communicable diseases (NCDs) are diet and lifestyle related diseases which not only predominant in developed countries but also in developing countries. The clinical expertise is required for blood pressure measurement and confirmation of diagnosis with investigations. So there is need for simple, easy, cost effective tool for screening large population for early identification of risk factors.
Objectives-• To study association of adiposity indicators with risk of hypertension
• To study association of various factors with risk of hypertension and develop a screening tool for identifying individuals at risk of hypertension
Materials and methods-A cross sectional study on 300 adults in age group (25-45yrs) from IT and Non IT sectors. Anthropometric measurements included weight, height, body mass index (BMI), waist circumference (WC), hip circumference (HC) with blood pressure levels measured using standard procedures. Percent body fat
(BF %), percent Visceral fat (VF %), was taken by OMRON(HBF 362) body fat analyzer. Results-Female participation was very low therefore analysis was done only for males.

∑ Prevalence of hypertension (130/85mmof Hg) was 11.3% & blood pressure increased significantly (p<0.05) with increase in adiposity.

∑ Compared to lowest tertile of BMI (<23) the odds ratio (OR) for risk of isolated systolic hypertension (ISH) in the highest tertile of BMI (>26) was (OR-2.174, 95%CI) while for highest tertile of VF i.e.>11% OR for risk of IDH was (OR- 2.34).

∑ Higher consumption of milk and milk products showed higher risk for IDH (OR 2.15) while for higher consumption of bakery products it was 1.53.

∑ Screening tool was based on significant risk factors showed subjects with higher score (>21) had higher risk for ISH, IDH, HTN (z value=3.43, 28.94, 1.83 respectively) compared to those with lower score (<17). The score grouping showed a linear trend and strong association with SBP and DBP (p-0.000 for both).

**Paper Name:** Differential Age Typology of Treatment Seeking Population with Substance Use at De addiction Centres across India

**Author:** Ashwani Mishra, Yatan Pal Singh Balhara, Swati Kedia, Sudhir K Khandelwal

**Abstract:**

Introduction
In recent years treatment seeking population at various de addiction centres has exhibited an upward trend. Hence it becomes imperative to understand the profile of such population, for better understanding the phenomenology of substance use. The present study aims at understanding the age typology of treatment seeking population with Substance Use at De addiction Centres across India.

Methods
The study is based upon the data obtained through Drug Abuse Monitoring System (DAMS). DAMS is the regular on-going activity at National Drug Dependence Treatment Centre (NDDTC), AIIMS, New Delhi, supported by Ministry of Health and Family Welfare. The DAMS provides socio-demographic information along with the current (last one month) and lifetime use of different substances like alcohol, heroin, opium, other opioids, cannabis, sedatives, cocaine, other stimulants, hallucinogens, volatile solvents and tobacco (yes/no) among treatment seeking population.
Results
The Mean (SD) for the total sample (n=1, 55,853) was 33.5(11.3). The absolute number of treatment seeking population over the period 2007-2015 increased substantially (2007-11063; 2008-11519; 2009-13283; 2010-12969; 2011-15966; 2012-20000; 2013-24157; 2014-29017; 2015-18152), thereby amounting to absolute percentage increase of 64%. The zone wise distribution across various parts of India was as- North-94084; South-29015; East-5719; North East-14623; and Central and West-12412. The mean age at the time of seeking treatment varied significantly (P<0.001) across different years and zone (minimum-north, maximum-south). Moreover, mean age of individuals reporting current use was significantly lower than those not reporting current use for any of the substances except alcohol and opium.

Conclusion
The DAMS data provides the evidence of increasing trend for substance use with varying age typology. Hence age appropriate treatment services should be developed across these centres. More studies are required to understand the reason behind the increase and differential age typology.

Paper Name: An evaluation of difference in prevalence estimates for factors associated with cardiovascular diseases using complete cases analysis and multiple imputation strategies

Author: Binukumar B

Abstract:
Background: Missing data leads to list-wise deletion often produces biased estimates and imputation of missing values from observed values may be an efficient approach to maintain the power of the study. Use of multiple imputation strategies overcome the situation of homogenous of observations, underestimates the variance in single imputation, and arrive at a valid statistical inference.

Objectives: This present communication aims to compare the prevalence estimates for factors associated with cardiovascular disease (CVD) by complete case analysis (CCA) and multiple imputation analysis (MIA) of Cardiovascular Risk Reduction in South Asia (CARRS).

Methods: We have used data from CARRS surveillance study, a survey conducted in a representative sample of 16,287 non-pregnant adults aged 20 and older from three mega-cities of South Asia (Chennai: 6906, Delhi: 5364 and Karachi: 4017). CCA include 10,976 participants and CVD risk factors such as tobacco use, hypertension, diabetes, overweight, central adiposity, elevated lipid levels and low intake of fruits and vegetables were considered. Multiple Imputation using chained equations (MICE) approach was adopted for imputing missing values. Models used were linear
regression and predictive mean matching for continuous variables and ordinal logistic regression for ordinal variables. Ten imputed datasets were generated. Diagnostics were performed on the imputed datasets and the convergence was also checked. Further we estimated the age and sex standardized prevalence of CVD risk factors using CCA and MIA.

Results: Although socio-demographics data are available for above 99% of the participants, there exists substantial missing information in anthropometric and laboratory investigations with maximum missingness for weight (23.1%). Crude prevalence of CVD risk factors in CCA and MIA had shown a difference of maximum 1.8% for central adiposity. Standardized prevalence in CCA and MIA standardized to regional populations reported a highest difference of 1.1% for tobacco smoking; whereas the same to the world population shown a difference of 1.2% for central adiposity. However difference observed for diabetes, hypertension, overweight/obesity and lipid parameters were very minimal.

Conclusion: Results of imputed data using MICE largely consistent with result of CCA. For epidemiological research with smaller sample size, use of multiple imputations may be substantial to increase the power.

Acknowledgement: B. Binukumar was supported by grant number 1 D43 HD065249 from the Fogarty International Center and the Eunice Kennedy Shriver National Institute Of Child Health & Human Development at the National Institutes of Health.

Paper Name: Twin Burden of Co-morbidity (TB with DM) and Awareness regarding Effect of Diabetes Mellitus (DM) on Treatment Outcome of Tuberculosis among Patients attending TB Clinic in a Tertiary Care Hospital in Delhi

Author: Vijay Tiwari, Piyush Verma, Sherin Raj

Abstract: According to WHO Report 2014, the estimated prevalence of TB in India is 2.5 million and 0.22 million deaths are occurring due to TB every year. An increasing prevalence of DM may counteract the positive effects of improved curative services for TB. In order to assess the co-morbidity of DM &TB and also awareness in patients regarding effect of DM on treatment outcome of TB, a study was conducted among 220 TB patients attending TB Clinic in Tertiary Hospital in Delhi. It was found that out of 220 TB patients, 36 patients (16%) were co-morbid. Among co-morbid patients, the age distribution skewed towards higher age groups as compared with only TB patients. Co-morbidity was higher among females (64%).
as compared to males (36%). According to Kuppuswamy’s Socio-economic status scale 61% all TB patients, 78% Co-morbid patients and 58% only TB patients belonged to Lower/upper lower socio-economic class. Majority of patients of all three categories were either illiterate or educated up to middle school. Co-morbid patients were found to be more physically activity than the only TB patients. Statistically significant association was observed in background variables (age groups, level of education, socio-economic status, residence and caste) with physical activity; dietary habits and tobacco/alcohol consumption etc. Statistically significant association in knowledge about DM with gender, level of education and place of residence was observed. Nearly 80% all aware cases and 50% of co-morbid cases were aware that high sugar intake, overweight or obesity are the leading risk factors for DM. Only one third of patients who were aware about DM, knew that diabetic patients are more prone to acquire TB infection. Less awareness was found regarding the effect of DM on the treatment of TB among all aware cases. In case of co-morbid cases, only 50% were aware that uncontrolled DM delays the cure of TB. Findings also suggested that the sputum conversion rate among the co-morbid patients was slow compared with the only TB patients. Statistically significant association was observed in background variables (age groups, level of education, socio-economic status, residence and caste) with knowledge regarding TB disease, (symptoms of TB, transmission of TB, diagnostic techniques for detection of TB disease, ways of preventing transmission, complications of incomplete TB treatment and comorbidity of TB with HIV). The study recommends that screening for DM in all patients diagnosed with TB, and screening for TB in people with DM should be made as a part of national programme strategy in RNTCP in accordance with the WHO Framework for collaboration of DM-TB. Also, counselling sessions should be made compulsory for TB & DM Co-morbid patients.

Session (3.25-4.25)

1. Medical Statistics 1

**Paper Name:** Evaluation of Performance of Hospital Service: A Graphical Approach  
**Author:** Dr. Chandra Tripathi  

**Abstract:**
Background: The objective of the study to evaluate the effective utilisation of hospital services and also assess their performance in a hospital using a graphical approach (Pabon Lasso Model) using combination of hospital indicators - bed turnover (BTO), bed occupancy rate (BOR) and average length of stay (ALS).

Method: This cross-sectional descriptive study was carried out in 2015 at the Institute of Human Behaviour and Allied Sciences, Delhi, India. This study involved various wards catering to psychiatric, neurological and neurosurgery facilities in the institute. Their performance was evaluated over 8 year period (2007-2014) using three performance indicators (BTO, BOR & ALS) to assess optimal utilisation of hospital facilities.

Results: Psychiatry department was initially located in quadrant IV in 2007 & shifted to quadrant III in 2014 which suggests department’s good quantitative performance with small proportion of unused beds. Similarly Neurosurgery department was in quadrant I at its inception in 2010, but shifted to quadrant III in 2014. However, Neurology department was located in quadrant III initially (2007), but shifted to quadrant II indicating either excess bed supply or less need for utilisation.

Conclusion: Pabon Lasso model can be used by hospital management to evaluate and improve the quality of health care services.

**Paper Name:** Statistical Analysis of Medical Data for Inventory Management in a Healthcare System

**Author:** Esha Saha

**Abstract:** Statistical analysis is a powerful technique in the field of healthcare that enables to draw a meaningful insights from a study in which medical data are collected through real-time observations, survey and from medical records. In the field of healthcare, there are several research problems ranging from clinical to operational, such as decisions on organ transplant, operating room planning, appointment scheduling, etc. One of such operational problem dealt in this paper is the inventory management in healthcare systems. Inventory management in healthcare is an upcoming area of research as efficiently managing inventory acts as a prerequisite for many planning and decision making processes in a healthcare system, such as forecasting and developing healthcare policy, design decisions related to location and storing, resource allocation, etc.

A study on healthcare is encouraged with the prime motive of enhancing the service level of patient. To meet the regular demand of the patient, the healthcare units need to be always prepared with the required resources so that no delay hinders the treatment of the patients who may arrive in any extreme health condition. Along with meeting regular demand, resources for handling emergency demand must also be
sufficient in the healthcare units to provide maximum service. To maintain all the necessary resources to fulfil the regular and unpredictable demand of patients, inventory of items required for patient treatment needs to be available when required at the right place and at the right time.

Inventory in healthcare systems includes that of healthcare goods i.e. drug products, single-use medical devices and healthcare equipment. The decisions regarding how much quantity of these healthcare goods need to be inventoried in the healthcare unit is of prime concern as there need to be a balance between the risk of being short of items and having excess of these items.

In course of solving such problem, there is a requirement to analyze various factors affecting the healthcare inventory problem, such as patient-related information (patient arrival and discharge rate, their length-of-stay, bed occupancy, etc) and operational factors (item usage quantity, their unit price, annual consumption, etc). If it is possible to accurately predict the factors-influenced demand pattern of healthcare inventory items then decisions regarding inventory management is reliable. Thus, this requires statistical analysis of the medical data of various factors influencing inventory management and for which the required medical data is collected from a hospital in eastern part of India. The medical data is summarized using numerical and graphical descriptive statistical methods. Inferential statistical procedure is applied to draw conclusions on the basis of sample data analysis.

Paper Name: Estimate the Concordance Correlation Coefficient between Automated Blood Pressure monitor and Mercury free LED Sphygmomanometer Device.

Author: Shashi Singh

Abstract:

Introduction: Automated Blood Pressure monitor is widely used to assess the blood pressure of the study subjects in community based research. This study examines the concordance between automated and mercury free LED sphygmomanometer devices. Methodology: This study has been carried out at Rural Health Training Centre, Ormanchi, Ranchi in August 2016. Thirty subjects aged between 40 and 70 years were enrolled and taken oral consent for this study. Left Arm Blood Pressure (BP) of the patients in sitting position was measured by automated device by three times in 5 min. intervals and readings were recorded by an observer (A) and again repeated assessment of same patients were done by mercury free LED sphygmomanometer Manual BP device by another observer (B) to avoid biasness. All readings of both devices were compiled in MS Excel and exported to SPSS for data analysis. Degree of agreement between two machines i.e., Automated (@OMRON Model no. HEM 7200) and mercury free LED sphygmomanometer
BP machine of different sets of readings were assessed by concordance correlation coefficient (CCC). Results and Conclusions will be shared during the presentation in the conference.

**Paper Name:** Longitudinal Model for Undifferentiated Hypovolemia Patient Data

**Author:** Mini Jayan, Marimuthu P, Binukumar B, Muralidhara K

**Abstract:**

**BACKGROUND**
Medical research often adopt the method of longitudinal studies where the association of patient's interventions and outcomes are assessed at multiple follow up times. The traditional statistical approach is to use repeated measures ANOVA (RM ANOVA) by treating the time points as categorical variable, whereas the mixed effect model treats time as a continuous variable. The later approach allows us to model a regression line for time, rather than estimate different means. However these data might not be complete for all the time points. In the mixed model approach, if a time point is missing, only the missing time point will be dropped unlike in multivariate approach where the subject will be dropped from the entire analysis.

**OBJECTIVE**
The objective of this study is to compare the estimates of RM ANOVA and mixed effect model.

**METHODOLOGY**
The data used here is obtained from patients with undifferentiated hypovolemia, a potential reversible condition which, if not treated early, results in multi organ failure, primary cause of death being rapid volume loss. Ultra-sonographic measurement of Inferior Vena Cava (IVC) diameter is one of the non-invasive modalities available to assess the volume status. Here IVC variability in patients are studied using the inspiratory (dIVCi) and expiratory (dIVCe) diameters of the IVC. Systolic, diastolic and mean arterial blood pressure are the predictor variables. All these variables were available at baseline and 5 other time points. Age, Gender and BMI are other available covariates which does not change across time (fixed covariates).

RM ANOVA was carried out for IVC variability to test the mean difference across different time points. Later a Multiple linear regression analysis was performed for the IVC variability with the fixed covariates. Further the use of a linear mixed effect model was explored by adding time dependent covariates to the model. Subsequent
comparison of estimates of these models were done.

RESULTS
Mauchly’s test in RM ANOVA showed that IVC variability does not satisfy sphericity assumption. Greehouse Geisser test gave a significant difference in IVC variability at different time points. A significant pairwise difference was shown at time points (3, 6) and (4, 6). The data turned out to be a peaked data with the peak at third time point. Marginal variance covariance matrix of the basic mixed effect model showed compound symmetry. Subsequent models showed difference in the structure of variance covariance matrix. P-values were obtained by likelihood ratio tests of different models.

CONCLUSION
Compared to RM ANOVA, mixed effect model gives a better understanding of the data. Since mixed models have a subject specific approach to analysis, a GEE model shall also be attempted to emphasize on group effects.

Paper Name: Modelling of Ordinal Outcome: Factors Associated with Extent of Nodal Involvement among Oral Cancer Patients

Author: Vishwajeet Singh, S.N. Dwivedi, S.V.S. Deo, M.A. Khan, V Sreenivas

Abstract:
Background:
The most common cancer among men in India is cancer of lip-oral. Understanding of burden of nodal involvement among them as well as its associated factors may provide clues to the clinician for better management of these patients. Hence, for effective prognosis of oral cancer, one also needs to understand etiology of nodal involvement.
Keeping in view of clinical relevance, the frequency of involved nodes is generally described in ordinal form, but no study have dealt with ordinal form of nodal involvement. As a first study in this regard, primary objective of this study was to find out the associated factors of nodal involvement.
Materials and Methods:
The data base on oral cancer patients available with the Dept. of Surgical Oncology, BRAIRCH, AIIMS, New Delhi, was used. Under this study, all histopathologically proven oral cancer patients who underwent for surgery including neck dissection were included. To find out the associated factors, nodes involved were grouped into four categories as, 0, 1, 2-4 and >4. Keeping in view of the related assumptions, the ordinal regression model namely partial proportional odds model was used for the analysis. For multivariable analysis, a subset of covariates was selected on its
clinical relevance and/or maximum p value as 25% under univariate analysis. The results were considered significant at 5% level of significant. Accordingly, the results in the form of odds ratio and corresponding 95% confidence interval (95% CI) were considered.

Results:
For multivariable analysis, gender, age, pain at presentation, sub mucous fibrosis, clinical T stage, skin involvement, clinical neck node, tumor growth type, tobacco chewing, smoking, alcohol, duration of symptom, duration of risk, oral site, tumor margin and Histopathological tumor size could be taken as covariates. The results related to variables following parallel line assumptions may be interpreted as per convention followed under usual analysis. However, those related to variables not following parallel assumption may be interpreted in view their patterns. Patient with pain at the time of presentation had higher chance of experiencing higher frequency of positive nodes [1.37 (1.06-1.78)]. Those who had presence of clinical node, were more likely to being the higher positive node group [2.47(1.77-3.44)]. Presence of SMF, was protective for higher frequency of node involvement [0.43(0.21-0.88)]. The positive tumor margin, is more likely for presence of positive node, but it was unlikely to have highest positive node group [2.17 (1.34-3.52), 1.54 (0.92-2.58) and 0.51 (0.15-1.67)]. However, as tumor size increases, the chance of being in higher positive node group is likely to increase [1.08 (0.87-1.34), 1.34 (1.05-1.71) and 1.71 (1.18-2.49)].

Conclusion:
Pain at time of presentation, presence of clinical node, positive tumor margin and increased size of tumor are most probable associated factors with higher nodal involvement.

2. Demography-2

Paper Name: Decelerating mortality at older ages and its future prospects

Author: Anurag Verma

Abstract: In this paper, we investigate the trends and future prospects of decelerating mortality at older ages in India and its states by age and sex. We observed and analyzed the trend and pattern of mortality by age and sex for the period 1971 to 2011, while the time series methods are used to make long-run forecasts, with confidence interval, of age-specific mortality in the India and its states from 2012 to 2061. In particular, we adopt Leslie matrix technique and Lee-carter model to project the population and age specific death rate, respectively. The
estimation of model parameters is based on the singular value decomposition technique, while the mortality index is forecasting using random walk with drift technique. Fitting of model has been shown graphically and the model found to be close fit. We have decomposed the change in deaths during 1971-2011 to understand the role of population growth, population ageing and death rates. In India, the trends in mortality is declining, while the level of mortality is not declining but rate of acceleration is declining and further in future expected to be decline. Also the death pyramid shows shift in mortality from childhood ages to adult and excess of adult mortality among male.

**Paper Name:** To assess the determinates of unmet need of family planning acceptance level, and trend of Uttar Pradesh

**Author:** Sarvesh Kumar, Neha Mishra

**Abstract:** The concept of unmet need points to the gap between women’s reproductive intensions and their contraceptive behaviour. In doing so it poses a challenge to the family planning programme to reach and serves the millions of women whose reproductive attitudes resemble those of contraceptive users but whom, for some reasons or combination of reasons, are not using contraception. By responding to the concerns of women with unmet need, programmes can serve more people and serve them better. Programmes can respond best if there is a strategy that focuses on women with unmet need as a distinct audience and clientele. Such strategies could reinforce other strategies. The present study is an application of different statistical models to study the family planning acceptance and unmet need in Uttar Pradesh using the India’s National Family Health Survey in the State. The Study has utilized the secondary data collected under the two rounds of National Family Health Survey conducted in the year 1998-99 and 2005-06 (NFHS-2 & NFHS-3) in Uttar Pradesh. It has tried to define the study variable “unmet need” under the multinomial regression and its possible determinants. Having estimated a class of multinomial multivariate logistic models, some determinants of unmet need have been studied. In a population like Uttar Pradesh where there is very high level of infant mortality and also the son preferences, an explicit attempt has been made to examine the effect of sex composition of children on unmet need of family planning. The result are significant in accessing not only the magnitude of the unmet need but also the reasons behind this unmet need at the individual, family and community level.
Paper Name: Model for Child Mortality Assessment under Varying Parity

Author: Arunabh Tripathi, G P Singh

Abstract: In this study a model is conceptualized to measure the child mortality under different parity of women such that a better strategy can be formulated to bring down child mortality rates. In the estimation of probability of child mortality some socio-demographic parameters are also taken in consideration. The estimates of parameters are obtained under Bayesian procedure. To obtain the Bayes estimates of parameters of the model a programme was written in the language of WinBUGS. Two different models are formulated for it and model fitting is observed by graphical approach along with the chi square test. It was found that Model –II fit better as compare to Model-I. The estimate of child mortality at parity 3, parity 4 and parity 5 were obtained as 0.06, 0.09 and 0.13 respectively from model-II. It was also found that Muslim women have higher chance of child mortality as compared to Hindu women in all parity.

Paper Name: Population Projection of India and its states using Bayesian Approach

Author: Abhinav Singh

Abstract: Projection of future population for a country by age and sex, are widely used for policy development, planning and research. They are mostly done deterministically, but there is a widespread need for probabilistic projections. In this paper we propose a Bayesian method for probabilistic population projections for country. The total fertility rate (TFR) and female and male life expectancies at birth are projected probabilistically using Bayesian models. The two models are then used in Bayesian framework for future fertility and mortality, which is combined using cohort-component method to obtain age-specific projection of the population by sex. The analysis has been made using Markov Chain Monte Carlo (MCMC) technique with the software OpenBUGS. Convergence diagnostics techniques available with the OpenBUGS software have been applied to ensure the convergence of the chains necessary for implementation of MCMC. The method is illustrated by making 40-year projection using Indian data for the period 1971-2011. The advantages of our method are: (1) it is projected two different demographic component separately; (2) it is based on historical data with no subjective inputs required and (3) it provides probabilistic point estimates of parameter as well as the projection along with highest posterior density (HPD) interval, that is derived from population number and vital
events, includes age specific death rates, life expectancies, age specific birth rate, total fertility rates and dependency ratios.

**Paper Name:** Knowledge, Attitude and Practice GAP in Family Planning Usage: An Analysis of Selected Cities of Uttar Pradesh

**Author:** Anjali Singh, Kaushalendra Kumar Singh

**Abstract:**

Background- The GAP between the knowledge of contraception and its actual practice is well recognized in the literature of family welfare studies. The present study assessed the relation between the level of knowledge and practice of contraception among the women and seeks to explore the reasons behind the Knowledge, Attitude and Practice - GAP (KAP GAP) regarding contraceptive uses in six cities of Uttar Pradesh.

Method- Present analysis was based on 17,643 currently married women aged 15 to 49. Bivariate analysis (χ² test) and multivariable logistic regression were performed for the analysis by taking KAP GAP as the dependent variable.

Result- The highest percentages of respondents (women) were in the age group 35-49 (40-45%) in all the districts considered. Knowledge of contraceptives was almost universal; female sterilization and pill were the commonly known methods. Information about the contraceptive methods was mostly obtained through husband. In the present study there was a highly significant association between age group, educational status of respondents, number of living children, wealth of the respondent, media exposure and husband’s education with the variable KAP GAP for all six cities. Health concern issue in all the districts was most prominent reasons for not using contraception.

Conclusion- The differences in the socioeconomic and demographic factor exits, which lead to KAP GAP in the family planning (FP) usages. Therefore, in designing effective family planning program, there is a need to understand the various factors which influence the practice of contraception.
Paper Name: Status of HIV infections among clinically suspected reproductive aged females attending S S Hospital, BHU.

Author: Ajay Singh

Abstract:
Introduction and Objective
Today, nearly half of all adults living with HIV around the world are women. Just under two-third of all people living with HIV are in sub-Saharan Africa, as are 76% of infected women. Among young people living with HIV in this region, around three out of four are female. Most HIV positive women have been infected with HIV through heterosexual sex. Information drawn from different studies shows that during heterosexual sex, women are about twice as likely to become infected with HIV from men as men are from women. Physically, women are more susceptible than men to HIV infection through heterosexual sex, and this fact alone means that special attention must be paid to protecting them if they are not to be disproportionately affected by the epidemic.
Epidemiological data world-wide shows that major mode of HIV transmission is sexual intercourse. The objectives of this paper are (i) to study the sero-positivity rate of HIV infections among reproductive aged females, and (ii) to find out the association of age, residential area, presence of migration, mode of transmission and different morbidities/symptoms with HIV status.

Data
This study is based on 2806 female subjects of reproductive age group referred for sero-positivity status at ICTC Department of Microbiology, IMS, BHU, Varanasi. All the samples of symptomatic and asymptomatic are double tested for HIV positivity by following strategy II/strategy III respectively (Elisa/Rapid/Retroscreen/ Retroquick) using WHO/NACO guidelines.

Result
$\chi^2 = 93.58$, $p < 0.001$ Among 2806 females cases, 341 (i.e. 11.8%) were HIV positive. The percentage distribution of HIV positivity was 3.56, 14.73 and 18.54 in the age groups 15-24, 25-34 and 35-49 years respectively. This shows that as the age increases, HIV positive cases also increase and there was significant difference in the positivity percentage of HIV and the age groups ($<0.001$). The percentage distribution of having experiences of multiple sexual contacts in HIV positive females is 60.34. The relative risk (R.R.) is 11.32 and 95% C.I. is 11.06-11.593. The positivity of HIV infection is also high (11.7%) and significant among women having the history of migration. The prevalence of infection is also high (17.33) in rural areas in
comparison to 5.36% in urban areas. The co-morbidities were absent in 37.35% cases whereas 62.65% were with one or more morbidities. The significant morbidities with HIV infection were fever, weight loss, weakness, loose motion, T.B., cough, anorexia and others including STD symptoms and AIDS symptoms.

Conclusions
These findings indicate a high HIV positivity prevalence among women in the reproductive age group (15-49 years). The analysis suggests that heterosexual intercourse is the major mode of transmission in women. Increased efforts to reduce the spread of HIV among this epidemiologically important group are urgently needed.

Paper Name: Community-based prevalence of genital Human Papilloma Virus (HPV) infection: A systematic review and meta-analysis

Author: Ravishankar N

Abstract:

Introduction
Cervical cancer probably represents the best studied human cancer caused by a viral infection and the causal association of this preventable cancer with Human Papilloma Virus (HPV) was well established. Worldwide there is a scarcity of data regarding the HPV prevalence and a vast difference exists in the HPV prevalence among the populations.

Objective
The aim of this meta-analysis was to determine the community-based HPV prevalence estimates in women from urban, rural set ups and women attending cancer screening clinics.

Methods
PubMed-Medline, CINNAHL, Scopus, and Google scholar were systematically searched for the studies that provided the prevalence of Human Papilloma Virus infection among asymptomatic women from 1986 till 2016 April.

Results
The final analysis included 32 studies comprising a population of 2, 24,320 asymptomatic women. The overall pooled HPV prevalence was 11% (95% CI: 9%-12%). The pooled HPV prevalence of 11% (95% CI: 9%-11%) was observed among women attending cervical cancer screening clinics. The pooled HPV prevalence was 10% (95% CI: 8%-12%) and 11% (95% CI: 4%-18%) from urban and rural areas
respectively, indicating higher infection rates among the rural women with the least access to cancer screening and cancer care.

Conclusion
The prevalence rates in this systematic quantitative review provide a reliable estimate of the burden of HPV infection among asymptomatic women from developed as well as developing nations. Policy makers should ensure equity in cancer prevention, care, and survival outcomes for the socioeconomically deprived populations from remote areas.

Paper Name: Modulation of metabolic profile with reduction in bone density during manifestation of arthritis: A correlative study
Author: Priyaranjan Dash

Abstract:
Objective
To investigate the relationship between metabolic markers and disease activity, parameters of bone metabolism and bone mineral density (BMD) at the lumbar spine (BMD-LS) and the femoral neck (BMD-FN) measured by dual X-ray absorptiometry in rheumatoid arthritis (RA) and Osteoarthritis (OA).

Methods
A number of two hundred and eighty seven (287) patients with final diagnoses of arthritis were enrolled in the study and classified into two study groups – Osteoarthritis (n=139) and Rheumatoid arthritis (n=148). The patient’s demographic and biochemical data including bone mineral density (BMD) values (performed by DEXA) were recorded at the Out Patients centers of Department of orthopaedics, Agartala Govt. Medical College, Agartala, Tripura. The data were collected in connection with the respective physician of the orthopaedics department. We have performed correlation matrix for all the data collected throughout the study.

Results
The demographic data were found to vary significantly between the osteoarthritis (OA) and rheumatoid arthritis (OA) study groups. Rheumatoid arthritis patients had a higher incidence of osteoporosis (65/139, 47%) than that in Osteoarthritis (50/148, 34%) healthy controls (5/30, 17%). They displayed lower BMD values than controls at positions of all detected regions. Minor correlations were observed between the serum metabolic markers and the BMD values but the correlation was found significant among the BMD values of AP spine, femur neck and the total body. The
data have shown a significant positive correlation between the BMD values of AP spine (L1-L4) and femur neck (both left and right).

**Conclusion**

These data suggest that, there is a firm association between the disease severity and inflammation (level of CRP, ESR and RF as inflammatory markers) in OA and RA patients. The minor correlations between the biochemical parameters including serum Ca, vitamin D, cholesterol and BMD value have also led to the conclusion that these markers may be a reflection of disease-activity-related systemic bone loss, and could be a predictor of BMD at the lumbar spine and femur in OA and RA. The positive correlations between the BMD values have suggested that both AP spine and femurs are equally affected during the pathogenesis of OA and RA.

**Paper Name:** Structural Equations Modelling Of Health Related Quality Of Life Of Patients With Lymphatic Filariasis

**Author:** K.T.Harichandrakumar & M.Kumaran

**Abstract:** Lymphatic filariasis (LF), commonly known as elephantiasis, is a neglected tropical disease. Infection occurs when filarial parasites are transmitted to humans through mosquitoes. Lymphatic filariasis is a major public health problem, 1.3 billion people are at risk of infection and about 128 million people are affected in 83 countries. Lymphatic flariasis is also a major public health problem in India. The disease manifestations of LF are acute and chronic in nature. The painful and profoundly disfiguring visible manifestations of the disease are lymphoedema, elephantiasis and scrotal swelling and it leads to permanent disability and thus irreversible. The acute episode of adenolymphangitis (ADL) among chronic patients was considered as co-morbidity. The LF patients are not only physically disabled, but suffer mental, social and financial losses contributing to stigma and poverty. The objectives of the study are to assess the level of impairments of HRQoL in LF patients when compared to normal subjects and to develop a Structural Equations Model (SEM) of HRQoL in patients with LF. HRQoL is a multidimensional construct which refers to a persons’ perceived physical, mental and social health. SEM is a powerful multivariate analysis technique. One important features of SEM is the ability to model constructs as latent (unobserved) variables using indicator (observed) variables and the estimation of latent variables from indicators allows for the quantification and separation of measurement error within the model. Furthermore, complex relationships between multiple constructs can be modelled with SEM. SEM consists of two models as Measurement Model and the Structural Model. Measurement model is an application of confirmatory factor analysis (CFA) and is used to examine the relationship between indicators and latent constructs. Structural Models utilize multiple regression paths among latent variables to test specific relationships between constructs.

The study was conducted in the LF endemic areas of Tamil Nadu and Pondicherry. The patients were randomly selected from the line listed cases of LF and the normal
subjects were also selected from the same locality for socio-economic and cultural matching. 239 filarial lymphoedema patients with different grades of lymphoedema and 204 matched normal subjects were participated in the study. The distribution of patients in relation to the clinical manifestations shows that, 5% had grade-1 lymphoedema, 40% had grade-2, 35% had grade-3 and 20% had grade 4 lymphoedema. The HRQoL was assessed by using SF-36 and it consists of 36 items. The overall HRQoL of patients with LF was found to be significantly lower (P<0.05) than the normal subjects. Among the lymphoedema patients, the HRQoL found to be decreasing while progressing the lymphoedema. The patients with recent ADL had reported more impairment in the HRQOL and the HRQoL is comparables between the genders. SEM of the HRQoL among the patients with LF was developed by considering the items and sub-domains of SF-36 as the indicators. The latents were derived from the indicators and the relationship of the latents with the indicators were developed using measurement model and also the relationship between the latents were explored using structural model. The relationship of the socio-demographic and the clinical characteristics with the HRQoL also explored using SEM. The different process involved in the development of SEM includes the model specification, identification, parameter estimation and testing the goodness of fit. The closeness of the variance-covariance matrix implied by the hypothetical model to the empirical data was evaluated by using chi-squared statistic and two goodness of- fit indices; the comparative fit index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The greatest advantage of SEM is the ability to manage measurement error, which is one of the greatest limitations of most of the models. SEM is an effective tool in clinical and epidemiological research to determine the contribution of each factor to the outcome by controlling other factors. The application of SEM has been used in many disciplines but it has yet to be extensively used in medical research and epidemiology. The details of SEM, the estimates of the parameters and the goodness of fit of the model will be presented and discussed in detail.

Paper Name: Bland-Altman analysis in agreement studies: When and how to use?
Author: Mahasampath Gowri. S, Antonisamy B
Abstract: Bland-Altman’s analysis or limits of agreement (LoA) is used to assess the relative agreement between two instruments or experimental methods where the continuous variables are measured in same scale, which is identical to a Tukey mean-difference plot. Over the past two decades (1986-2014) the method was cited in almost 20,000 publications according to web of science, an online subscription-based scientific citation indexing service maintained by Thomson Reuters.

Two important assumptions of the method are (i) distribution of normality for the difference between the methods to be evaluated (ii) constant change in difference
i.e. the difference does not depend on magnitude of the methods. In some scenario the assumptions are violated, but still the method is used to explain the systematic differences in the instruments, which is in turn results in faulty and unreliable conclusion. To overcome these issues, different methods to construct the plots are suggested. This paper is intend to review those techniques and to provide a tutorial, when and how to use the plot effectively and meaningfully. The LoA construction with conservative approach, non-parametric approach and regression approach were compared and explained in detail with the datasets.

In case of violation of normality assumption, a non-parametric LoA is constructed and interpreted, explained in detail using the data of two methods measuring the WBC counts in Blood. The use of regression approach to construct the LoA , when the distribution of the difference shows heterogeneous pattern i.e. not constant over the magnitude of measurement is discussed in detail using the data measuring postoperative cylinder values and the cylinder values calculated using the formula to predict the cylinder values in the eyes. Additionally, the LoA calculation to evaluate (i) the agreement between repeated measures data and (ii) the agreement method to validate a standard against new technique is explained.

The bland-Altman analysis or LoA is very simple and direct method which can be done with basic statistical knowledge when the assumptions are satisfied. The interpretation of the method is very meaningful and easy. The complexity arises due to the violation of assumption which can be solved and used further.

**Session (4.35-5.35)**

1. Medical Statistics-2

**Paper Name:** A comparison of approaches to assess change in EEG phase synchronization on epilepsy patients

**Author:** Ravi GS

**Abstract:** Cortical neuronal synchronization has been implicated to be an important mechanism driving the occurrence of epileptiform activity during sleep. The cortical phase synchronization index (SI) values measured prior, during and post an Inter-ictal Epileptiform Discharges (IED) in various sleep stages yields correlated
data. Such data are characterized by correlations among repeatedly measured values of a variable for an individual. The variability between and within subjects and IED should be treated separately while analyzing such correlated data. Mixed Effect Models (Random effect models) and Generalized Estimating Equations (GEE) are the techniques which are known to offer more flexibility in handling correlation among repeated observations and other covariates.

In this study, an attempt is made to compare different approaches in answering questions about within and inter-individual differences in changes over time. A data on SI among 40 Juvenile Myoclonic Epilepsy (JME) patients (20 drug naïve and 20 patients on drug) was used for the study. The ensemble SI was calculated, for delta (δ; 0.5-3Hz), theta (θ; 4-7Hz), alpha (α; 8-12Hz) and beta (β; 13-30Hz) frequency bands during video EEG monitoring of the patients. The changes in SI’s over these stages were assessed using different approaches for repeated measures data. The sleep stage at which an IED occurred and the drug group were treated as fixed factors and the inter-IED variability as random. The effects of other covariates like, gender, age of the patient and sleep quality were also assessed.

There were 108 and 125 IEDs among on drug (OD) and drug naïve (DN) group respectively divided among various sleep stages. The results show that that, group on antiepileptic drug group have significantly higher SI in delta (OD: 0.541 ± 0.035, DN: 0.534 ± 0.032, P=0.003) and beta bands (OD: 0.45 ± 0.008, DN: 0.445 ± 0.014, P=0.001), and lower SI in alpha (OD: 0.5 ± 0.02, DN: 0.511 ± 0.029, P=0.002) and theta bands (OD: 0.496 ± 0.015, DN: 0.505 ± 0.025, P=0.002). The random variation between baseline IED values (intercepts) was found to be significant in all the bands. The GEE approach is also applied along with the random intercept and random slope models. They offer more flexibility in handling categorical predictors and time dependent covariates. The detailed results would be presented.

**Paper Name:** A comparative study among PLHIV and Non PLHIV in India  
**Author:** Nitesh Adichwal  
**Abstract:** In December 2002, WHO and Joint United Nations Programme on HIV/AIDS (UNAIDS) estimated that 42 million people worldwide were living with HIV/AIDS. Since the beginning of the epidemic, about 30 million people have died of AIDS; the AIDS deaths in 2002 alone were 3.1 million. In 2002, 5 million adults and children were newly infected with HIV. Of these, more than 95 per cent were living in the developing world. The overall HIV prevalence in India has been estimated to be approximately less than 1 per cent and India continues to be in the category of low prevalence countries. As India is a vast populous country, the low prevalence of HIV
do not reflect so much problem but in reality the number of cases reported were quite large. In the year 2002, an estimated 4.58 million people were living with HIV infection in India, the second-highest number in the world, after South Africa (NACO, 2002).

Because of the heterogeneity between states, the HIV estimates for India show a variation with respect to certain background characteristics like variation among age group, caste, region, marital status etc. Therefore, in order to get complete picture of the scenario, it is important to study these estimates with respect to background characteristics. There are many studies which have discussed about knowledge, perception, attitude and risk behaviour of HIV/AIDS among men and women but these studies have not dealt with the differential among PLHIV and non-PLHIV men and women and the opportunistic infections which can be caused primarily because of HIV. Due to lack of knowledge, people have misconceptions about HIV/AIDS and related infections. So it is important to know their perception and attitudes towards HIV/AIDS so as to provide them proper knowledge about AIDS and HIV transmission through programmatic intervention. This study makes an attempt to know how the knowledge, perception and attitude changes among PLHIV (cases) and non-PLHIV (control) people with respect to certain background characteristics and that in conformity with gender wise variation. It also tries to know the risk behaviour and HIV related infections such as TB and STI among them.

**Paper Name:** A Comparison of Performances of Machine Learning Methods for Predicting Outcomes after Traumatic Brain Injury (TBI)

**Author:** Vineet Kamal, Ravindra Mohan Pandey

**Abstract:** In biological or medical sciences, due to recent advancement in data-mining and predictive modelling techniques, one can highlight and chose any predictive or classification model depends upon its application, utility, and context. In this study, comparison of performances of predictive models based on machine learning methods [logistic regression (LR), classification and regression tree (CART), and artificial neural network (ANN)] in terms of overall performances (Brier score), discrimination ability [Area under ROC curve (AUC)], and calibration ability (Hosmer-lemeshow test’s p-value) to predict in-hospital mortality and unfavourable functional outcome at six months post admission in patients with moderate or severe traumatic brain injury have been done in both development and external validation dataset. Predictor variables were demographic [age (years), sex], clinical [motor score, pupillary reactivity, limb movement, cause of injury, major extracranial injury, duration before admission (hrs)], secondary insult (hypotension), CT reports (degree of MLS, SDH, EDH, basal cistern effaced, tSAH/IVH). In external validation dataset, for LR, CART, ANN, area under AUC (95% CI) were 0.86 (0.81, 0.90), 0.80 (0.75, 0.84), 0.79 (0.74, 0.83) respectively.
0.86), and 0.86 (0.82, 0.91), respectively to predict in-hospital mortality; and 0.91 (0.87, 0.94), 0.87 (0.83, 0.91) and 0.91 (0.87, 0.94), respectively to predict unfavourable outcome at six months. Calibration was good for all methods (Hosmer-Lemeshow test’s p-value>0.05) for both outcomes except CART model for in-hospital mortality (H-L test p=0.005). Brier score was least by ANN method, followed by LR, and CART method for both outcomes. More or less similar performances were seen for LR and ANN, but these both methods outperformed CART to predict outcomes in terms of each respect of performance measures.

**Paper Name:** Trail of Labour versus Repeat Caesarean Section for the Women with a Previous Caesarean Section

**Author:** B. S. Dhillon, Nomita Chandhiok

**Abstract:**

Objective: Objective of current study was to study the outcome of trial of vaginal birth after Previous Caesarean Section (PCS) and indications for repeat caesarean section at teaching hospitals in India.

Methods: Prospective data was recorded on management practices, associated complications and mortality for a period of 8 months in 2005-2006 at 30 medical colleges/teaching hospitals for delivery.

Results: A total of 155863 deliveries occurred during the study duration, there were 28.1% (n=43824) caesarean section and (10.1%) (n=15664) were the number of previous caesarean section. In 84% (n=13151) had repeat caesarean delivery and 2513 (16%) delivered vaginally. A trial of labour was planned in 4035 (25.8%) women. The success rate of VBAC was 62.3% with 2513 women had successful vaginal delivery and 1522 (37.7%) delivered by repeat caesarean section. Major indication of repeat caesarean section was CPD (52.9%), foetal distress (25.8%), severe PIH/eclampsia (5.0%), previous 2 CS (0.7%), APH (1.4%) and others (2.7%). In majority, surgical technique was conventional and in 3.7% the Misgav-Ladach technique was used. Scar dehiscence and surgical complications were observed in 5.4% and 4.0% of cases respectively. Blood transfusion was given in 7.0% and post-operative complications were seen in 6.8%. Perinatal and maternal mortality was 18.0/1000 and 257/100000 deliveries respectively.

Conclusions: Safety in childbirth for women with prior caesarean is a major public health concern. Repeat caesarean section and planned vaginal birth after caesarean section are both associated with benefits and harms and correct management represents one of the most significant and challenging issues in obstetric practice.
**Abstract:** Statistics plays a vital role in the research of medical sciences. It helps medical researchers to design studies, data collection, analyze data from medical experiments, help interpret the results of the analyses and collaborate in writing articles to describe the results of the study. It is the theoretical science of the inferential process, especially the planning and analysis of experiments, surveys, and observational studies. In research related with medical field, the data are qualitative and quantitative both, so that the researchers use statistical tests and to determine results from experiment. Now a days researcher also use statistical softwares for analyzing their data.

This particular article discuss about some statistical techniques, tools and brief idea about the related statistical softwares.

---

2. Demography-3

**Paper Name:** Cause Specific Premature Mortality in India

**Author:** Durgesh Shukla, RavindraNath Mishra

**Abstract:**

Introduction: Premature mortality is a measure of unfulfilled life expectancy that estimates the average years a person would have lived additional after the premature death. Premature mortality rate gives more weight to the death of younger people of which many preventable. Measuring premature mortality, rather than overall mortality, underlies the intent of the country health ranking to focus attention on death that could be prevented. The measures of mortality like potential year of life lost are good indicators to comparing the relative importance of different cause of death for a population. The objective of the present analysis is to compare premature mortalities for 8 leading causes of deaths in India.

Material & Methods: Secondary data from the Sample Registration System (report on cause of death: MCCD 2013) and Census 2011 of India is used for analysis. Year
of Potential life lost (YPLL) was calculated following its standard procedure and had been used as index to estimate the premature mortality.

Results: Over all YPLL due to 8 leading causes of deaths in India was 10793596 i.e. 964 years per lac population. Among these, the circulatory diseases ranked the first by life lost 210 years per lac population followed by certain infectious and parasite diseases (204 years per lac population). YPLL due to injury, poisoning & other consequences and the diseases not classified were almost same and were 164 and 153 years per lac populations respectively. The lowest YPLL was found by diabetes mellitus (19 per lac population); while diseases of neoplasm, digestive system and respiratory system observed YPLL below 100 “between” 55 to 85 per lac populations.

Conclusion: YPLL due to only 4 diseases e.g. circulatory, certain infectious and parasite, injury, poisoning & other consequences and not classified were contributing almost 75% of YPLL; the government has to focus on these to reduce the premature deaths as many premature deaths can be averted through the life style modification.

**Paper Name:** Female Education and Fertility in Slums of Varanasi  
**Author:** Payal Singh, RavindraNath Mishra  
**Abstract:**

Introduction: Human fertility is the actual reproductive performance of a woman; conventionally defined as the number of live births in the entire reproductive life time. Excessive replacement, especially now days in poor, is creating socio-economic burden to their own family as well as the nation, thereby it’s a major concern. Slums are the poor and less educated and less caring to children that lead to continued reproduction. However, education in these poor plays its role to limit the family size; hence the objective is framed as below:

Objective: To assess the effect of female education on fertility.  
Material & methods: 600 eligible couples were randomly selected from 5 randomly selected slums following PPS sampling. The woman’s age, history of live births, desire of children, contraception practices and the education were elicited from women respondent. The fertility measures taken were average born and proportion proceeding to more than 3 live births. The analysis was carried only on those who had a minimum of 5 years marital duration. The average born children by education for each marital duration group was computed and finally logistic regression was carried to identify the effect of education to reproduce more than 2 live births by logistic regression.
3 live births. Compared to mothers of education more than intermediate the risk to precede 3 or more live births was nearly 69%, 65% and 74% less among primary or below, middle and intermediate level ≥10 years marital duration had given ≥ 20 years of marital duration respectively. About 75% women with ≥ Results: The average children born were decreasing as the education of mother increasing; from 2.31 in illiterates to 1.76 among above intermediate in 5-9 marital durations and from 4.11 in illiterates to 3.17 among above intermediate in

Conclusion: Education of girls has to be promoted for substantial reduction in fertility.

**Paper Name:** The Level and Determinants of Excess Fertility in Uttar Pradesh

**Author:** Ruchi Mishra

**Abstract:** This paper focuses on the level and determinants of excess fertility by taking the difference in realized fertility and desired fertility for women of Uttar Pradesh. The study analyzes the effects of some potential socio-demographic factors such as place of residence, religion, years of schooling, wealth index, age at marriage and current age of the women in the study on the gap between actual and desired fertility level. The data used for this study are obtained from Measurement, Learning and Evaluation Project for The Urban Health Reproductive Health Initiative, Uttar Pradesh. For analysis purpose, only those females are considered who have completed their family and stated that they want no more children. Logistic regression is used as a statistical tool to examine the effect of various socio-demographic variables on the level of excess fertility. The findings of the study support that the females who have high level of education, belong to rich wealth index, dwelling in non-slum areas, and have higher age at marriage generally have desire for less number of children and lower level of realized fertility as compared to the females living in slum areas, belonging to poor wealth index, having low level of education and are married at an early age. The inverse relationship between the years of schooling and desired fertility as well as actual fertility may play the vital role in framing social policies to reduce fertility level. A similar relationship is found between age at marriage and actual fertility level. Generally, the main factor for keeping fertility at a high level is the low educational attainment of females.

**Paper Name:** Modelling the sex ratio at birth under sex preferred stopping rules

**Author:** Anjali Pandey, Kaushalendra Kumar Singh, Anjali Singh
Abstract: Being a crucial demographic indicator, secondary sex ratio or sex ratio at birth has been extensively studied through several empirical and theoretical researches. Son preference has been considered as one among the various biological, socio cultural and behavioural factors influencing the sex ratio at birth. In South and East Asian countries, predominantly in India, China, Singapore, Taiwan, Hong Kong and South Korea son preference and sex selective abortions have brought upon a noticeable difference in the number of males and females being born (number of male births being greater than number of female births). Various socio cultural and religious practices are believed to be the basis of these sex preferences in such societies. This leads to a natural curiosity that whether the sex ratio at birth will be affected by son preferred stopping rules of the couples in such societies? Almost all the studies conducted in this direction assumed that despite the fact that the couple may have sex preferences, the sex of the child being born is purely random and the couples do not have any control over it. However, the availability of sex selective induced abortion has interfered with this aspect and together with sex preferred stopping rules it may influence the sex ratio at birth. (Yadava, 2013) investigated this issue by incorporating the possibilities of sex selective induced abortion and considered three stopping rules, each of these stopping rule being male preferred, and two having provision for sex selective abortion. Therefore, the couples may choose to abort the child if it's not of their preferred sex. (Yadava, 2013) They derived the expression for probability of having k children, expected number of children and sex ratio at birth under these three stopping rules. Working on same stopping rules, we derived the different expressions for probability of having k children, expected number of children and sex ratio at birth under the same three stopping rules.

This study utilizes mathematical modelling approach to estimate the sex ratio at birth with help of more realistic assumptions and varying the values taken by the parameters involved in the models. In the societies where sex selective induced abortions have been prevalent, change in the sex ratio at birth has been a matter of great concern for the policy planners. But assessment of such underplaying phenomenon is difficult due to unavailability of data. Modelling helps in these scenarios by providing an alternative approach in the absence of empirical support. The current paper implies that male child preference together with sex preferred induced abortions will lean the sex ratio at birth towards male child, in particular for smaller values of n.

Paper Name: Differential and Determinants of Neonatal and Postnatal Mortality in India

Author: Akash Mishra, Payal Singh, Ravindra Nath Mishra
Abstract:

Introduction: Fourth MGD goal was framed to focus on reduction of Child Mortality. India achieved much reduction but could not achieve the fixed target of 27. High IMR is due to high neonatal deaths; about 2/3rd total infant deaths. For quick reduction in IMR, determinants of neonatal and postnatal deaths need to be focussed separately.

Objectives: To assess the differential pattern of neonatal and postnatal mortality

Data & Methodology: Data of NFHS-3 was used. A total 2876 deaths were under five and 2508 were of infant. Ratio of neonatal and postnatal deaths was 2.2:1. Variation of neonatal and postnatal mortality was analysed using logistic regression.

Observations & Results: Compared to mother’s education higher secondary or above, the risks of neonatal mortality was 2.53 times in mothers either with no education or primary and 1.90 times higher in secondary level while risks of postnatal mortality were 2.57 and 2.20 times higher in no education and primary mothers respectively. The risk of neonatal mortality was highest among born to poorest families compared to richest families. The risks of neonatal mortality to born of 2nd or 3rd order and 4th to 5th were lesser by 30% and 22% compared to born of 6th order; while 1st order born had almost same risk while risk of postnatal mortality was lesser by 29% only in born of 2nd or 3rd order. Neonatal mortality of male children was higher than female, but postnatal mortality was almost similar. The size of child at birth had influence to both neonatal as well as postnatal mortalities. The risk of neonatal mortality was almost similar in children delivered at home or institution, but postnatal mortality was 1.37 times higher among children delivered at home.<20 and 20-40 years were 1.47 and 1.18 times higher; while postnatal mortality was 1.50 times higher among <Results: Place of residence and religion did not contribute either neonatal or postnatal mortalities. The risks of neonatal mortality was 1.89, 1.41 and 1.32 times higher in Central, East and North regions compared to south; while postnatal mortality was higher by 1.46 times in Central region only. The risks of neonatal mortality among mothers

Conclusion: The findings suggest that rising marriage age, limited births through promoting family planning devices and MCH care can bring a quick reduction to both neonatal and postnatal mortality and thereby in IMR and the target fixed can be achieved well in time.

3. Epidemiology-3
Symptomatic morbidities in localities around solar evaporation ponds and Union Carbide Factory Bhopal.

Author: S. Singh, N. Banerjee

Abstract:

Background: World biggest man made industrial disaster took place in UCIL Bhopal in 3rd Dec. 1984 which affected the environment and ecology of Bhopal by polluting air, water, and soil for long term duration. After 31 years, there is need for establishing cause and effect relationship to study the effect of this environment factor on health. MIC gas were affected more than 5.74 lakhs population. Residual chemicals left over in plant and in solar evaporation ponds would percolate down to ground water and on consumption this toxic water cause morbidities among the population living around factory complex.

Objective: Symptomatic morbidity status among the people living around the solar evaporation ponds and Union Carbide Factory in reference to the consumption of alleged contaminated ground water.

Methods: There was a study planned to establish the cause and effect relationship between allegedly present toxicants in water and the observed symptomatic morbidities. Tools like secondary data review on water analysis and cross sectional morbidities survey of people living in 14 localities around the solar evaporation pond and behind the Union Carbide factory were used for this purpose. Matching control localities were selected at about 10 kms away.

Result: Symptomatic morbidities survey did reveal statistically significant higher level ($p>0.01$) of symptomatic morbidities among the resident living in affected localities in comparison to that of localities in control area.

Conclusions: It can be concluded that the higher symptomatic morbidities could be due to earlier toxic gas exposure, bad environmental/sanitation conditions. There is need for clinco– epidemiological study of the population.
Background: cancer has become one of the ten leading causes of death in India. Breast cancer is the most common diagnosed malignancy in India, it ranks second to cervical cancer. The increase in overall worldwide burden of breast cancer may be attributable to increasing life expectancy and widespread adoption of westernized life style. Breast cancer has large number of etiological factors which will be increase the risk, an increased risk of breast cancer in women with a family history of breast cancer has been demonstrated by many studies.

Objective: The objective of the present communication is to drive the pooled estimates of odds ratio and 95% CI of family history and breast cancer.

Material and Method: Thirty seven case-control studies published during the period 2000-2016 were retrieved through Pub med, Google search, and previous reviews were included in the meta-analysis. Positive family history was taken as risk category. DerSimonian-Laird method to compute pooled odds ratios (ORs) and confidence intervals (CIs) and the Mantel-Haenszel test to assess association between family history and breast cancer.

Results: The result of thirty case-control studies pooled to determine the association between family history and risk of breast cancer. The pooled odds ratio along with 95% CI of fixed effect model was 1.51(1.38-1.64) and pooled estimate of random effect model was 1.64(1.45-1.84) Test of Heterogeneity Q= 54.27 P=0.026 statistically significant and I2 was 33.7%.

Conclusion: A patient family history may aid clinical diagnosis and contribute to disease risk assessment and prediction. Therefore it is possible that identifying people with a family history of disease could act as an additional motivator for them to change their lifestyle or participate in disease screening.

Paper Name: Morbidity Profile of Truck Drivers: A South Indian Study
Author: Xavier ChristuRajan V, Anil C Mathew, Ramamoorthy V, Dhanush M, PremAnand ES.

Abstract:

Introduction: Recent studies in the other part of the country have shown that workers on the transport industry are more predisposed to cardiovascular morbidity and other health morbidities due to their improper diet and sedentary life style. The commonly seen morbidities reported among truck drivers were Diabetes Mellitus, Obesity, Hypertension and heart disease. To the extent of our knowledge, not many studies were done in India to assess the morbidity profile of long distance truck drivers. This study was aimed to assess the common morbidity profile among the truck drivers.
and associated demographic and clinical parameters considered relevant to the development of Cardio Vascular Disease.

Materials and Methods: This was a cross sectional study conducted among 175 Truck drivers aged 18 – 60 years at the Walayar check post, Coimbatore, Southern India. After getting informed consent, a detailed physical examination was conducted. Socio demographic questions were also included. Descriptive statistics were calculated. Poisson regression model was performed to determine the independent predictors of the number of morbidities of the truck drivers include diabetes mellitus, overweight and obesity raised systolic Blood Pressure (BP) and raised diastolic BP.

Results: It was observed that 50% of them were overweight. Hypertensives were 41.1% and diabetics were 12%. About 27.4% had raised systolic BP and 29.1% had raised diastolic BP. The other common morbidities were low back ache (34.3%), hip pain (24.6%), knee pain (9.7%), shoulder pain (5.4%), diminished vision (4.6%), giddiness (2.3%) and hearing defects 1.1%. Age was a significant independent predictor [OR 1.04 (95%CI 1.03-1.05)] for the total number of morbidities.

Conclusion: We observed a higher morbidity among the truck drivers. It is observed that more periodic health appraisals are needed for drivers at older age. Truck drivers contribute to the country’s economy, and their morbidities should be addressed.

Paper Name: Prevalence and Associated Factors of Diabetes among Truck Drivers
Author: Dhanush M, PremAnandES , Xavier ChristuRajan V, Ramamoorthy V , Anil C Mathew

Abstract:

BACKGROUND Diabetes mellitus is an increasing problem worldwide and among those groups at high risk are truck drivers, which is mostly attributed to their lifestyle and work. There are several risk factors that predispose these truck drivers to type 2 diabetes such as lack of exercise, unhealthy meal choices and obesity. It is thus important to determine the prevalence of diabetes among truck drivers and the common risk factors in order to target health policies and programmes towards decreasing their risk of acquiring type 2 diabetes.

MATERIALS AND METHODS This cross-sectional study was conducted at the Walayar check post in South India. 175 truck drivers who passed through this point were recruited in the study after obtaining their informed consent.
RESULTS It was observed that 21 (12%) truck drivers were diabetic. The factors that positively correlated with the risk of diabetes were older age (Odds Ratio (OR) = 1.126, P < 0.05) and moderate alcohol consumption (OR = 0.229, p < 0.05) were found to be associated with a lower risk of diabetes mellitus.

CONCLUSION we observed a high prevalence of diabetes among truck drivers. Truck drivers contribute to the country’s economy and their morbidity should thus be addressed. It is imperative to design appropriate road and job policies for truck drivers and to target them for periodic health evaluation for early detection of diabetes.

KEYWORDS: Truck drivers; Diabetes; Physical activity; Older age.

Paper Name: Joint analysis of cognitive and circadian variation in schizophrenia and Bipolar I Disorder

Author: Thomas Pramod, He Fanyin, Mazumdar Sati, Wood Joel, Bhatia Triptish, Gur Ruben, Gur Raquel, Buysse Dan, Monk Tim, NimgaonkarVishwajit, Deshpande Smita

Abstract:

Background: Impairment in cognitive variables and alterations in circadian function have been documented among patients with schizophrenia (SZ) and bipolar I disorder (BP1), but it is not known whether joint analysis of these variables can define clinically relevant sub-groups in either disorder.

Objectives: To evaluate the pattern and relationship of cognitive and circadian function in SZ and BP1 patients with respect to diagnosis and indices of clinical severity.

Methods: Among patients with SZ and BP1, cognitive function was evaluated using the Penn Computerized Neurocognitive Battery and circadian function was assessed using the Composite Scale of Morningness/ Eveningness (CSM). Clinical severity was estimated using the Global Assessment of Function (GAF) scale, and age at onset of illness (AAO). The patients were compared with community based non-psychotic control individuals and non-psychotic first degree relatives of the SZ patients. The cluster distributions of cognitive function, circadian function and clinical severity were investigated and identified clusters compared across diagnostic groups.

Results: When all the participants were analyzed, the cognitive domains could be separated into two clusters. Cluster 1 included the majority of control individuals and non-psychotic relatives, while SZ patients predominated in Cluster 2. BP1 patients
were distributed across both clusters. The clusters could be differentiated by GAF scores, but not AAO. CSM scores were not significantly correlated with individual cognitive domains or with the clusters.

Conclusions: Clusters based on levels of cognitive function can discriminate SZ patients from control individuals, but not BP1 patients. CSM scores do not contribute to such discrimination.

December 2
Session (10.25-11.45)

1. Biostatistics 1

**Paper Name:** Evaluation of biomarker in case of heterogeneous data through Rayleigh Mixture ROC Curve  
**Author:** Azhar Uddin

**Abstract:** The Rayleigh Mixture distribution can be applied when the data is heterogeneous or population consists of subpopulations. We have proposed Rayleigh Mixture Receiver Operating Characteristic (ROC) model and discussed its properties. Optimal cut off value of biomarker using Rayleigh Mixture ROC model is also obtained. Area under the Rayleigh Mixture ROC Curve (AUC) and its variance also derived. Estimates of the parameters of Rayleigh Mixture ROC Curve are derived by using the Maximum Likelihood Method (MLE) via EM algorithm. The proposed model is validated by using the simulation studies and real life example of head trauma data. It is found that Rayleigh Mixture ROC Curve is better than Bi-Rayleigh ROC Curve when data is heterogeneous.

**Paper Name:** Long-term outcomes of glaucoma drainage devices for glaucomapost-vitreoretinal surgery with silicone oil insertion: a prospective evaluation. 
**Author:** Kulwant Kapoor  
**Abstract:** Purpose To evaluate long-term success of the Ahmed glaucoma valve (AGV) for refractory glaucoma after vitreoretinal surgery with silicone oil insertion. Methods Prospective non-comparative evaluation of patients who underwent AGV
insertion for management of postvitreoretinal surgery glaucoma, post-silicone oil removal. Intraocular pressure (IOP), visual acuity, and glaucomatous neuropathy status were evaluated preoperatively and at multiple follow-up visits postoperatively. Success, using Kaplan–Meier analysis, was determined at the 12-month follow-up visit and at the last follow-up. Factors associated with failure were analysed. Results Twenty-seven eyes of 27 patients with a mean age of 28.3 ± 15.2 years underwent a superior AGV implantation. The average follow-up after AGV implantation was 17.11 ± 8.36 months (range: 9–60 months). Kaplan–Meier survival analysis revealed a 62 % success at 12 months and 37 % at 5 years. A 48 % rate of complications was noted, 22 % of which were vision-threatening. Factors analysed, including patient age, interval between vitreoretinal surgery and silicone oil removal, interval between vitreoretinal surgery and AGV implantation, and phakic status, were not found to be associated with higher failure rates. Conclusion Long-term success of AGV implantation for glaucoma after vitreoretinal surgery with silicone oil insertion is better than that reported for trabeculectomy, though complication rates remain high. Keywords Glaucoma .Silicone oil .Ahmed glaucoma valve .Vitreoretinal surgery. Refractory glaucoma

Paper Name: Latent Class Analysis for building sub-groups based on susceptibility to adverse health outcomes related to cardiac events
Author: Ankita Dey, Sugata Sen Roy, Diganta Mukherjee

Abstract: Latent class analysis (LCA) is considered to be an equivalent methodology for Factor Analysis, typically used for dichotomous or polytomous variables. LCA can be used in health research for identifying disease sub-types or diagnostic categories and building typologies in medical conditions. Diagnosis of disease is based on the traditional method of assessment of the value of diagnostic indicators such as medical signs, symptoms and medical tests. To achieve this objective, the evaluation of sensitivity and specificity (i.e. the probability that a person is positive on the indicator when the disease is present and the probability that a person is negative on the indicator when the disease is absent respectively) is the common practice. But it requires the knowledge of the presence or absence of the disease. Latent class analysis provides a tool for correct diagnosis without the prior knowledge of whether the patient is suffering from the disease or not. The traditional Latent Class models assume the underlying scheme as simple random sampling from a large population. The two population parameters of interest in a typical problem of LCA are the unobserved proportion or size of the latent classes and the conditional probabilities of an item response given the membership in a latent class. The assumption of simple random sampling scheme is violated in most of the real life situations where the data are actually collected under more complicated and hierarchical structure of data collection phases. Under these circumstances the usual estimates of the population parameters account for biases in the standard error. In the present study, attempts have been made to describe a model considering the
usual violation of the sampling assumption. Several indicators each having binary response options have been used to divide the population into two latent classes. Estimation procedure has been discussed in the context of maximum likelihood method. This study demonstrates the use of Latent Class Analysis to segregate the population according to the susceptibility of individuals to adverse health outcomes in two latent classes based on their observed measures on the risk factors as a method of medical diagnosis.

**Paper Name:** Comparison of Probit and Logistic Regression Models in the Multivariate Analysis of Dichotomous Outcomes: A Simulation Study

**Author:** Amrutha Jose, Dr.Mariyamma Philip, Dr.Manjula M, Dr.Thennarasu K

**Abstract:**

PURPOSE: Probit and logistic regression models are members of the family of generalized linear models, used for estimating the functional relationship between dichotomous dependent and independent variables. The current study is to find performance of logistic and probit regression models under different conditions.

OBJECTIVE: To compare the performance of probit and logistic regression models under multivariate normal.

METHODOLOGY: A simulation study was conducted in which an artificial data of a dependent and three independent variables were randomly generated using multivariate normal distribution with appropriate mean vector. Various conditions were imposed in terms of variance-covariance matrices assuming correlation (high, moderate, low & no) between dependent and both continuous and categorical independent variables. The dependent variable was transformed into dichotomous for different percentages of occurrences of outcome: 5, 10, 15, 20, 30, 40, 50, 65 & 90. In order to examine the effect of sample size the data was generated under different sample sizes: 30, 60, 90, 120, 200, 500 & 1000. For each of the combinations, 1000 simulations were carried out. Probit and logistic regression analyses were performed and the parameter, standard error and probability estimations were obtained. The likelihood ratio test was performed and root mean square errors (RMSEs) were estimated. Null and residual deviances, different pseudo R2 measures, AIC, BIC and correct percent prediction (CPP) were computed and the models were compared.

RESULTS: AIC of logit (0.5: 1391.852, 0.1: 640.590 & 0.9: 643.958) and probit (0.5: 1391.852, 0.1: 640.620 & 0.9: 643.953) showed higher values when the proportion of outcome, p=0.5 and decreased towards both lower (p=0.1) and higher (p=0.9) proportions. Similar trend was also seen in BIC. The RMSEs of logit and probit models were same (0.5: 0.500, 0.1: 0.307 & 0.9: 0.302) showed decreasing trend to both extremes. It is evident from the study that larger sample size gives higher AIC and are similar for both models (1000: 1391.852). For smaller sample size, it was
higher in logit (30: 45.052) compared to probit (30: 45.040). BIC also showed similar
trend. The RMSEs of logit (30: 0.300) and probit (30: 0.302) showed lower values for
higher correlations. RMSEs gradually increase with increase in sample size (1000:
0.348) and was similar in both the models. CPP of logit (1000; 0.1: 91.63) and probit
(1000; 0.1: 91.59) showed that as the sample size increases with decrease in
proportion of outcome, CPP also increases. In lower correlation with smaller sample
size, the prediction was higher in logit model (30; 0.1: 67.04) than probit model (30;
0.1: 54.55).
CONCLUSION: RMSEs showed that logit model fits better for lower sample sizes
(up to 200). Larger the sample size, probit and logit models fit equally well shows
smaller RMSEs especially.

Paper Name: Which is Preferred Measure of Heterogeneity in Meta-analysis and
Why?
Author: Mona Pathak  S.N. Dwivedi  S.V.S. Deo
V Sreenivas  Bhaskar Thakur

Abstract:
Background: The statistical heterogeneity can be interpreted as the excess of
between study variations over within study variance. Heterogeneity assessment is
unavoidable under meta-analysis as it guides the path to choose appropriate
synthesizing method. Its consideration further strengthens the drawn inferences. In
the literature, there are various measures of heterogeneity, such as, Q statistics, H2
statistics, H2m statistics, R2 statistics and I2 statistics. The present work mainly
aimed to appraise these methods and find the preferred measure of heterogeneity.
Material and methods: To begin with, for appraisal of above methods, a systematic
review was performed to assess the effectiveness of neoadjuvant chemotherapy
using PubMed and Cochrane databases. A total of 17 studies were found to be
eligible comparing at least one of the considered outcomes out of 1239 identified
records. The major outcomes were overall survival (OS), disease free survival
(DFS), recurrence free survival (RFS), time to Loco-regional recurrence (LRR), time
to distal recurrence (DR) and breast conserving surgery (BCS). As obvious,
exploration involving multiple outcomes was expected to provide better supporting
evidence. Various heterogeneity measures were compared on the basis of
theoretical criteria as well as their analytical results on the real data available from
eligible studies. Theoretical criteria include: (1) it should depend on the extent of
heterogeneity. (2) It should be scale invariant. (3) It should not rely on number of
studies and (4) it should have easier interpretation. For appraisal, 95% bootstrap
confidence interval was calculated for each of the considered outcomes. To compare
the precision, a unit free measure as ratio of standard error and point estimate of
heterogeneity, was developed and used.
Result: Out of 17 eligible studies, OS, DFS, RFS, LRR and DR were reported by 14, 6, 12, 10 and 12 studies respectively. In case of each heterogeneity measure, studies were considered heterogeneous if 95% bootstrap CI was not including the null value of that measure. Q statistics showed presence of heterogeneity for all of the outcomes except overall survival \[ p = 0.567, 95\% \text{ CI (0.069, 1.065)} \]. All other measures consistently showed mixed results, homogeneity of studies in the case of OS, DFS, and heterogeneity for DR. However, for RFS, only I\(^2\) showed heterogeneity \[ 39.307 (3.581, 75.034) \]. The precision assessed by the unit free measure was similar for all heterogeneity measures on observed data. However, I\(^2\) and R\(^2\) satisfy each of the theoretical criteria but I\(^2\) has more appealing and easier interpretation.

Conclusion: I\(^2\) remains a preferred measure on the basis of theoretical comparison as well as its performance on observed data.

**Paper Name:** Relative Risk estimates from regression model when the proportion of outcome is high

**Authors:** Jothilakshmi D, Sebastian George, Jeyaseelan Lakshmanan and Visalakshi Jeyaseelan

**Abstract:**

When the outcome variable of interest is dichotomous, a logistic regression model, which directly yields an estimated odds ratios adjusted for the effect of covariates. However the probability of outcome is high, the odds ratio and relative risk are different. Hence we recommended log-binomial model. However, Convergence problems may arise with binomial regression models. Therefore Poisson regression with a robust error variance will be applied.

**Methods and Materials:**

Binary response variable \( Y \) depends on a set of \( k \) explanatory variables, \( X=(X_1, X_2, X_k) \), which models the log odds of probability of “success” as a function of explanatory variables.

\[
\log[\pi(x_i)] = \alpha + \beta x_i.
\]

The India SAFE study was conducted during the period April 1998 to September 1999 by the Indian Clinical Epidemiology Network (IndiaCLEN) in collaboration with the International Clinical Epidemiology Network (INCLEN) as part of the world studies of Abuse in the Family Environment (WorldSAFE) study. This data was taken Vellore site.
Results:

In the domestic violence study, women who were harassed was very high. Poisson regression with robust error variance provided an appropriate model and it gave relative risk (RR) and confidence interval for RR for alcohol consumption to be 1.105 (1.033,1.183) and husbands who are illiterate was found to have a risk of 1.002(0.945,1.062).

Conclusion: Poisson regression with a robust error variance model gives better estimates.

2. Demography-4

**Paper Name:** Ranking and Classification of Districts of Bihar in light of NFHS-4  
**Author:** Alok Rahul Singh

**Abstract:** In the present study, an attempt has been made to rank and classify districts of Bihar on the basis of important health related indicators having sound theoretical reasoning and availability of data in NFHS-4 report. The selected indicators have been categorised into three dimensions, which are demographic features, cognizance and family health. Data procured on all the indicators have been converted into common scale of 0 to 1 and then dimensional index for each dimension has been computed. Each of the dimensional index provides a basis for comparison of districts within the dimension. The overall index has been computed by combining all three dimensional indices. This overall index provides the basis for ranking and classification of districts as a proxy to overall health. The ranked series of districts has been divided into three groups. First having overall index below lower quartile termed as under developed districts, these districts need more attention of the administration. Second having overall index more than upper quartile termed as developed districts and third of remaining districts termed as developing districts. This ranking will provide administration a basis for planning equitable overall health development.

**Paper Name:** Determinants of low birth spacing in India: A Cross-sectional study  
**Author:** Parul Puri

**Abstract:**
Background:
The adequate spacing of two subsequent births is very important for both maternal and child health. It can work as a natural birth control measure for the reduction of fertility level of the nation. The prime objective of the study is to explore the determinants of low birth spacing (LBS) in India.

Methods:
Third round of National family health survey (NFHS-3, India, 2005-06) data were used. Bivariate & multivariate logistic regression model was applied to identify the possible determinants of birth spacing with 95% confidence intervals of the estimated odds ratios.

Results:
Analysis suggests that low birth spacing (LBS) is substantially less among higher educated parents. The likelihood of LBS was found to be lowest among the respondents belonging to richest wealth quantile (AOR=0.655, CI of AOR: 0.587-0.730) and 8.8% lower amongst women who do some work (AOR=0.912, CI of AOR: 0.868-0.959). As compared to respondents who had never heard of family planning, the likelihood of LBS was 6.7% less in those who were partially (AOR=0.933) aware about family planning and 27.6% lesser among those who were fully (AOR=0.724, CI of AOR: 0.658-0.797) aware about family planning.

Conclusion:
The health effects of birth spacing remains relatively unnoticed despite a frequent urge to space pregnancies in order to improve maternal and child health. Efforts should be made to encourage couples for modern contraceptive use, higher education, and exclusive breastfeeding. As we know that our country could not achieve the targets fixed by National Population Policy (NPP) & Millennium Development Goal (MDG) regarding infant, child and maternal mortality in a fixed level till now. In order to achieve these targets; programs and policies must assess ways to provide equitable access to reproductive health interventions to mothers before or soon after delivering a child.

**Paper Name:** A Community based study on Quality of Life (QOL) and its associated factors using WHOQOL-BREF among Elderly Population in Varanasi, Uttar Pradesh, India

**Author:** Neha Seth

**Abstract:**
Background: Quality of life (QOL) among elderly is a deserted and overlooked issue especially in developing countries including India.

Aim: To assess the QOL and its associated factors among elderly population.

Materials and Methods: A community based cross-sectional study was conducted of elderly people living in the five colonies (Dashaswamedh, Bhalupur, Saket Nagar, Sundarpur, Samneghat) of Varanasi city, Uttar Pradesh, India. Data on QOL assessed by World Health Organization Quality of Life BREF (WHOQOL-BREF) questionnaire. The results were expressed in terms of mean and SE of mean. Student T tests and one way ANOVA were applied to compare the mean scores of different variables under the four domains.

Results: A total of 223 elderly people were included in this study. Out of 223 elderly people 139 (62.3%) were Males and remaining were females. The mean age of the study population was found to be 68.45 ± 7.58 years. The mean QOL score for all the elderly persons put together was 79.70 ± 14.73, indicating that on an average, the population as a whole had moderate to good quality of life. The highest score was for the environmental domain with mean 24.98 and standard deviation of 6.14 and the lowest was for social relationship domain with mean score of 9.68 and standard deviation 2.3.

Paper Name: Quality of Age Data: A Comparison between Two Recent Indian Censuses
Author: Minakshi Vishwakarma, Akhilesh Yadav, Prof.Chander Shekhar

Abstract: In India Census has started in late 19th Century. The first synchronous census was started from 1981. The Indian Census is the most credible source of information on Demography, Economic Activity, Literacy and Education, Housing & Household Amenities, Urbanisation, Fertility and Mortality, Scheduled Castes and Scheduled Tribes, Language, Religion, Migration, Disability and many other socio-cultural and demographic data since 1872. The age is an important demographic characteristic of any population provided by census or any other survey. Usually age data is used to describe the age structure of population and to forecast population growth. It provides valuable information for planning and formulation of various policies and programs which are based on age for example different pension program for old age population, child development program for children age less than 5, Dhanlaxmi Yojna for survival of girl child. But the quality of age data in census or sample survey is an issue it is affected by the common tendency of human which is digit preference. Human tends to round number to the nearest 5 or 10. This is also known as age heaping when people round their age to the nearest 5 or 10. In this paper we tried to explore the age heaping of census data and also the trend of
improvement over the period of 10 year. To measure this we use Wipple index and Mayer’s Index. In the current census i.e. 2011 the quality of age data comes in rough category by Wipple index. Age heaping is more among Male as compare to female also it is more among rural population as compare to urban population. From 2001 census improvement in quality of age has been observed in 2011 census. In 2001 census Wipple index falls in very rough category. While in current census first time date of birth of individual has been asked, also the education level has been improved in one decade. The same observations are come in to the picture with Mayer’s blended index. We also tried to observe the variation in age heaping among few states of India. In case of Kerala the Wipple index falls in the category of average quality while Bihar falls in rough category. The quality of census data regarding demographic indicator age is not so good and this is present among all states and strata of population. There is significant improvement in it is prominent over the period of time.

Paper Name: Assessment of Predictors of Neonatal Mortality in India
Author: Apurba Shil, ParulPuri

Abstract:

Background:
The increased reach of health programs in India, during the past few decades, has significantly contributed to decline in postnatal mortality including infant, child & other mortality. However, reduction in neonatal mortality remained negligible. About 70% neonatal deaths take place within a week after birth.

Objective:
The prime objective of this study is to explore the socioeconomic & demographic determinants of neonatal mortality.

Methods:
3rd round of National Family Health Survey (NFHS-3), India, 2005-06 data was used for this study. Bivariate analysis followed by multivariate analysis (Binary logistic regression) was performed to identify the determinants of neonatal mortality.

Results:
The odds of neonatal death was nearly 1.8 times higher (OR=1.79; 95% CI: 1.33-2.42) among births to Central region than the Southern region India. The estimated odds of neonatal death was 2.75 times higher (OR=2.75; 95% CI: 2.21-3.43) among
smaller size neonates than the average & larger size neonates. The odds of neonatal death was found 2.39 times, 2.73 times & 1.84 times higher among those mothers who had illiterate, primary & secondary level of education as compared to mothers who had higher education. The risk of neonatal death among 1st (OR=1.43) and more than 4th (OR=1.35) order births was positively higher than 2nd to 3rd order births. Although it was found that the odds of neonatal death was 1.46 times higher among adolescents (OR=1.46; 95% CI=1.06-2.01) mothers than old age mothers.

Conclusion:
Based on the findings, the study proposes various approaches to address the increasing contribution of neonatal deaths in India. The findings suggest that raising mother education & providing healthy diet to pregnant women, can bring a quick substantial reduction of neonatal mortality and the target fixed can be achieved. So, it is necessary to encourage & support women for higher education. This study proposes to encourage the combination of the continuum care for maternal, neonatal, and child health by integrating a family-community based service delivery approach with the existing health care system.

Keywords: Neonatal mortality, birth order, odds ratio, multivariate, Binary logistic regression.

**Paper Name:** Probability model approach for infant and child mortality: An evidence EAG states of India

**Author:** Mritunjay P. Singh, D. Kumar and R. D. Singh

**Abstract:** Infant and Child Mortality which are the two dimensions of mortality below five year age, and they are highly related to parity. In which infants have higher risk of mortality related to parity and so as in the children below five year age. it is evident that defining the direct measurement of that phenomenon is very tedious job. So we made to indirect approach of measurement with use of probability models. In this study we try to establish a model by which we can get the estimates of infant and child mortality separately at a time with the help of National Family Health Survey- III (2005-06) data of Empowered Action Group(EAG) states of India.

3. Health Statistics 1

**Paper Name:** Disparity in Terms of Haemoglobin Level in Females of Uttar Pradesh
Author: Tapan Kumar Roy, Brijesh Singh, Om Prakash Singh

Abstract: Anaemia continues to be a major public health problem in India and in spite of several plans and programs launched by government of India, there is no significant decline in the prevalence of anaemia and still a large proportion of women tend to anaemic. Anaemia refers to low level of haemoglobin in the blood. Haemoglobin transfers oxygen from lungs to other tissues and organs from of the body. It is one of the main causes of morbidity, mortality in reproductive age and a key factor to low birth weight these days. Due to poverty, inadequate diet, pregnancy, lactation, poor educational level and poor access to health services women become an easy prey for anaemia. The present study has used NFHS-III Uttar Pradesh data and tried to explore the differentials of anaemia among females of reproductive age group (15-49 years) in Uttar Pradesh. The available evidence clearly shows that anaemia is very widely prevalent in the reproductive age of women. However, the degree of severity varies across socio-economic, demographic and biological factors.

Paper Name: Primary Immunization Coverage & Socio - Demographic Factors Related To It among Children in Rural Area of Lucknow- A Cross Sectional Study
Author: Vijay Singh, Rahul Katiyar, Naim Ahmed, Jai Vir Singh

Abstract:

Context: India has one of the lowest immunization rates worldwide despite a longstanding Universal Immunization Program (UIP) that provides safe, effective and free childhood vaccines. India has about 61% immunization coverage and the situation is even worse in rural areas. Various factors are responsible for this low immunization coverage in India.
Purpose: This study aimed to evaluate Primary Immunization Coverage & to establish various socio-demographic factors related with it.
Methodology: It was a community based cross sectional study. A total of 210 children were selected using cluster sampling technique from rural area of Lucknow. The mothers were interviewed about the immunization status of their 12-23 months old children. A pre-tested semi-structured questionnaire was used for data collection.
Results: Out of 210 children that were taken for the study, 58% were completely immunized while 42% were partially immunized or not immunized (incomplete immunization). Vaccination status of female children was better (71%) than the males (43%). Immunization among children belonging to unreserved category was significantly higher (64.3%) than children of reserved categories (56.5%).
Conclusions: Almost 30 years have passed since beginning of UIP (1985) still this part of the country is lacking behind the national average of primary immunization hence sustained efforts are required to achieve universal coverage of immunization.
**Paper Name:** QALY: A MEASURING TOOL  
**Author:** Pradeep Jangid  

**Abstract:** QALY (Quality Adjusted Life Year) is a generic term used to measure the effects of health/medical interventions provided or offered in the treatment of diseases of interest, in terms of both quantity and quality of life. It is the arithmetic product of life expectancy and a measure of the quality of the remaining life-years. When combined with the cost of treatment involved in treatment, cost-utility ratios emerge. The use of QALYs in resource allocation decisions does mean that choices between patient groups competing for medical care are made explicit. Health planners increasingly have to prioritise their expenditure because they have limited sources to meet the demands that are placed on healthcare services. QALYs and cost-utility analysis provide additional information and are another piece in the complex jigsaw puzzle—the health service—which they are endeavouring to solve. However, QALYs are far from perfect as a measure of outcome due to number of methodological and technical shortcomings.

---

**Paper Name:** Statistical tool in the effect of Rajyoga Meditation on Quality of Life  
**Author:** Dr. Vijay Babar  

**Abstract:**

**Introduction**  
“Positive psychology,” is a recent method which is being employed in psychological treatment protocols, in an effort to better understanding about happiness, meaning of life, character strengths, and how all these can be developed and enhanced in life to have a better quality of life]. Positive emotions promote discovery of novel and creative actions, ideas and social bonds, which in turn, build an individual’s personal resources; ranging from physical and intellectual resources, to social and psychological resources. Positive emotions play a crucial role in enhancing resources and coping in the face of negative events.

**Objectives:** Psychological studies have shown that brief period of mindfulness meditation significantly improves critical cognitive skills. But, there are no studies which have assessed the effects of Brahma Kumaris Rajayoga Meditation (BKRM) practice on positive thinking and happiness in life. The present study was designed to test the hypothesis is BKRM enhances positive thinking and that essential to attain
higher levels of self-satisfaction and happiness in life.

Material and Methods: This study is a cross sectional comparative study which was done between Rajayoga meditators and non-meditators. This study was conducted at BKRM Centres at Wardha, Maharashtra State, India. Two Hundred subjects were selected for this study, which included those practising BKRM in their normal routine life (n=100) and non-meditators (n=100) who were aged 40.95 ± 14.23 years. Self-reported Oxford happiness questionnaire (OHQ) was administered to all subjects and their happiness scores and status were assessed and compared. Items related to self-satisfaction in life were selected from the OHQ and compared between meditators and non-meditators. Participants completed self-reported OHQ, from which data of happiness status and self-satisfaction in relation to meditation duration and frequency were analyzed by descriptive statistics and test of hypothesis.

Results: Mean happiness scores of BKRM were significantly higher (p<0.001) in meditators as compared to those in non-meditators. The number of meditators experiencing happiness status were significantly higher (p<0.05) in comparison with non-meditators. Additionally, meditators scored significantly higher on self-satisfaction items (p<0.001) as compared to non-meditators. There was no correlation between age and years of meditation practice with happiness score and self-satisfaction score.

Conclusion: BKRM helps in significantly increasing self-satisfaction and happiness in life by enhancing positive thinking. Irrespective of age and years of short-term or long-term meditation practice, enhanced positive thinking increases self-satisfaction and happiness in life.

**Paper Name:** Status of Adolescent health in Girls of Uttar Pradesh  
**Author:** Chandra Pandey, Anjali Yadav, Priyanka Sahu

**Abstract:**

Background: India constitutes of about 125 crores people. Adolescents constitute one fifth population of India in the age group 10 to 19 years. Out of 25 crores of the adolescents about 11.3 crores are girls. There is very little data available on health status of these girls particularly about physical body changes, puberty, menarche and other growth related challenges. This calls for attention to study these dimensions of adolescent health.

Objective: The objective of this study is to describe the health status of the adolescent girls in the state of Uttar Pradesh.

Method: The data for this study has been taken from Concurrent assessment of
health and family welfare programs and technical support to districts of Uttar Pradesh, a state wise study conducted by Department of Health and Family welfare, Government of Uttar Pradesh with technical support from European commission. This study was conducted in 2005-06 and collected information from all 71 districts of the state covering all dimensions of maternal and reproductive health including adolescent health. The present analysis is based on 1775 adolescents girls aged 10 to 19 years taken from 1378 households of three districts of Uttar Pradesh namely Rae Bareli, Sultanpur and Faizabad. Results: To assess the problem, girls were divided into two broad groups. In this 59.1% girls were in 10-14 age group and 40.9% were in 15-19 age group. We have focused on two main problems of girls, menstruation and RTI. Out of 1775 girls, 42.8% experienced menarche. Mean age of the subjects at menarche was 13.64±1.19years. Out of those who achieved menarche about 13.4% girls had problems related to menstruation. The most common menstrual problems are painful periods by 58.8% and short periods experienced by 18.6%. Out of which only 3.3% girls go for the treatments. For the consultation 60%, 8%, 4%, 28% girls have gone to the Private Hospitals, Government hospitals, Home Remedy and Others respectively. RTI related problems are vaginal discharge by 2.5%, pain in lower abdomen by 6.2% and lower backache by 4.1 percent. Out of which 18.2% girls go for the treatments. Three fourth of girls were gone to Government hospital and one fourth girls were gone to Private hospital for the treatment. Most (83.7%) girls used tidy clothes during menstruation and many girls (96.7%) did not go for the consultation. Menstruation can be improved by awareness. Perception of symptoms ‘being it normal’, ‘treatment being expensive’, ‘feeling shy’, ‘lack of female health workers’ were identified as major barriers for non-seeking treatment of RTIs. Improving literacy and increasing awareness level about sexual and reproductive health is required to reduce the incidence of RTIs among rural population.

Paper Name: Levels, Differential and Factors Associated with Falls among Older Adults in a Tertiary Care Hospital of Delhi: A Cross-Sectional Study

Author: Abha Agarwal

Abstract:

Background & Objective: Falls among older adult population (≥ 60 years) are a public health problem. A fall can result in lasting and critical consequences; including injuries, long term disability and dependency. Considering the increasing ageing population and socio economic changes in India, measure to keep older people healthy and active are the utmost importance. This paper aim to presents the
magnitude of the fall among the older adult (≥ 60 years), its differential and factors associated with falls in hospital setting. Method & Material: WHO definition of falls was adopted wherein a fall is defined as an event which results in a person coming to rest inadvertently on the ground.

A sample of 850 older adults was included from the Geriatric OPD, AIIMS, New Delhi.

A systematic random sampling was adopted to select the subjects from OPD, every fifth patient attending the OPD was interviewed for data collection after taking the informed consent. Reference Period for data collection for fall was previous one year. The data was collected from February 2016 to May 2016.

Results: The level of falls in Indian older adults was found to be 22% with SE of 1.5% and 95% CI as (19%, 25%). Out of 168 subjects with fall, 117(69.6%) had fall once, 43(25.6%) had fallen twice and 8(4.8%) had fallen for three times and above in past in one year. The mean age was 68.6 yrs in falls while 65.3yrs in no fall group. The significant difference was observed between the mean age of two groups, with fall and no fall (t= 6.33, p=0.000).The percent of fall in the age group 60 to 64 years was 27.3 %( n=46), 28.5% (n= 48) in age group 65-69yrs and 46.4% (n=74) in the age group 70 and above. Overall an increasing trend of fall with age is observed.

The unadjusted logistic regression was applied to identify the risk the risk of fall in age group 65-69yrs (OR= 2.02, p= 0.002, 95%CI (1.29, 3.15) when compared with the age group 60-64yrs. Similarly the OR for fall in 69+ is 4.13(p=0.000, 95%CI = 2.70, 6.30) more when compared with 60-64yrs. No significant gender difference was found between the two groups (p=0.291), indicating no relationship of fall by sex. The other predictor variables BMI, socio economic status, frailty, medication and chronic disease were also found to be significantly related with fall (p<0.05).

The adjusted logistic regression was applied to see the joint effect of the predictor variables age, BMI, sex, socio economic status, chronic disease, medication and frailty. The predictor variables age, BMI and frailty have significantly contributed in predicting the fall (p<0.05).

The risk of fall among underweight is higher (OR=2.46, p=0.040; 95% CI = 1.04, 2.56) as compared to those who were from normal category. Similarly, the risk of fall among over weight is also higher than those who were from normal category (OR=1.74; p= 0.006**; 95% CI 1.17, 2.57). The risk of fall is also significantly higher (1.25 times) in frail cases as compared to non-frail subjects (OR=1.25; p= 0.014; 95% CI=1.21, 5.21).

Interpretation & conclusion: The identified risk factors from the study were age, BMI and frailty. Findings of the study suggested that large scale studies need to be conducted at community level to develop the strategies for prevention of falls among older adults for quality life.
Session (11.50-1.05)

1. Biostatistics 2

**Paper Name:** Computation of Conditional Power - A case study of oncology phase II trial  
**Author:** Rajneesh Singh

**Abstract:** Objective of this paper is to explain the conditional power (CP) through a case study that is single arm, two stages, oncology phase –II with sample size of 63 cancer patients. The study treatment claimed 0.2 probability of surviving after 2 years. There are 23 patients have been enrolled in stage – I and followed for a year. On the basis of stage – I study data; statistician will make a decision whether study will continue in stage – II for remaining patients 40 patients i.e. 63-23 = 40 (63 patients are sample size in which 23 patients already enrolled and follow up for an year in stage - I ).

A SAS macro is developed to generate customized reports for conditional power. The calculation of conditional power (CP) is based on Z statistic and KM survival plot. A confident result is obtained from case study for continuation in stage –II. A conditional power (CP) from stage – I data is playing a very vital role to evaluate the decision for continuation of study in next stage i.e. stage - II. However apart from conditional power (CP), the clinical investigator decision on safety issues and their statistical proof provide the essential evidence to make a conclusion that whether study will carry on to the next stage or it will be stopped.

Keyword: Conditional Power (CP); Kaplan Meier (KM)

**Paper Name:** Statistical Methods to Analyse Ordinal Categorical Data Arising From the Clinical Trial of Drugs from the Pharmaceutical Industry  
**Author:** Nazneen Shariff

**Abstract:** The pharmaceutical industry gives rise to square tables of ordinal categorical data from the clinical trial of newly manufactured drugs. The Data arising from such trial have three dimensions: Drug X pre-treatment score X post- treatment score .The Proportional Odds Model is assumed to fit the data. The model is initiated by an underlying grouped continuous unobservable random variable. The ordinal categories on the post-treatment score is taken as a grouped continuous random
variable. The data is stratified according to the pre-treatment score for purpose of homogeneity. Statistical tests of the null hypothesis are formulated through the parameters of the proportional odds model. Three methods for testing the null hypothesis are examined: Regular Wilcoxon test, the classical Analysis of Covariance (ANACOVA) and a novel procedure existing in literature: Wilcoxon Van Elteren test for stratified data. A simulation study was carried out to examine the existing methods with the new method of application. The three methods are evaluated on a real data set of square ordinal responses from the clinical trial of a new drug. Theoretical results pertaining to the new Wilcoxon Van Elteren method are derived and verified on the real data set.

Conclusions:
The following conclusions can be drawn from the study of the statistical methods on square tables of ordinal data:
The three methods can be regarded as suitable aids in drawing inferences from square tables of ordinal data through the parameters of the proportional odds model. The Wilcoxon Van Elteren test is proving to be most efficient in the presence of a pre-and post-treatment relationship when compared with the regular Wilcoxon test and the ANACOVA.
The ANACOVA appears to be robust in that it gives the correct nominal significance level. But it loses power relative to the Wilcoxon Van Elteren test when there is a strong relationship between the pre-and post-treatment score.
In analyzing ordinal data from a drug induced symptom change study, the post-treatment score can be identified as a sensitive variable in detecting treatment differences.
Theoretical calculations on the parameter ‘q’ give consistency to the application of the Wilcoxon Van Elteren test.
The Proportional Odds Model is a satisfactory fit to the square ordinal data.

Paper Name: Contamination adjusted intention to treat analysis - Using instrumental variables to adjust for treatment contamination in randomised controlled trials

Author: Irene Elizabeth Joy, Sumithra Selvan, Johnson Pradeep, Krishnamachari Srinivasan, Tinku Thomas

Abstract:
BACKGROUND & OBJECTIVES:
Treatment contamination occurs when study participants do not receive the treatment to which they were randomised. Treatment contamination can occur through treatment non-adherence (not receiving the recommended intervention because of treatment intolerance or patient preference) and treatment crossover
(receiving the intervention intended for the other group in a trial) which can produce misleading findings. Just as non-adherence is common in clinical practice, treatment contamination in randomised controlled trials is not a small or infrequent problem. Contamination adjusted intention to treat (CAITT) analysis complements the intention to treat approach by improving the accuracy in estimating the size of treatment benefit for a patient who receives the treatment. The exact effect size of a treatment is relevant whenever you weigh a treatment against negative consequences like side effects, against treatment costs in cost effectiveness analyses, or against another treatment. This method is demonstrated using data from a cluster randomized trial.

METHODS:
Six villages from rural Bangalore were randomized to either community health worker supported enhanced care or usual care. A total of 260 adult depressed women formed the final participants for the analysis. The outcome measures were number of women who sought and completed treatment, number of clinic visits, duration of treatment with antidepressant, changes in severity of depression (HDRS) and changes in quality of life [WHO-QOL (Brev) scale].

RESULTS:
Adjustments were made in the statistical tests to account for the clustering effect of village for the primary outcomes. In this study, the number of clinic visits and weeks of treatment on antidepressant medication were significantly greater in the TI group compared to TAU group treatment even after performing intention-to-treat analysis as well as adjusted for clustering effect in the analysis. Using CAITT analysis, the effect of treatment assignment on outcome observed (number of clinic visits and weeks of treatment) was adjusted by the percentage of assigned participants who completed the treatment. Treatment completed was considered as an instrumental variable (if at all, the study groups were contaminated, we expected the percentage of patients who completed the treatment to be greater in control group also).

INTERPRETATION & CONCLUSIONS:
Intention to treat analysis estimates the effect of recommending a treatment to study participants, not the effect of receiving the treatment. CAITT can complement the intention to treat approach by producing a better estimate of the benefits and harms of receiving a treatment. It uses the statistical technique of instrumental variable analysis to address contamination. However, contamination is not always just a nuisance factor—it can demonstrate important factors such as how well patients tolerate treatment side effects.

Paper Name: Statistical Method for analyzing Microarray Gene Expression Data in Identifying Differentially Expressed Genes
Authors: Shanmugasundaram Devika, L. Jeyaseelan, Sebastian George

Abstract:
Microarray experiment allows us to study the expression profile of hundreds to thousands of genes simultaneously. These expressions could be from treated samples and the healthy controls. One of the main goals in the microarray experiment is to find out the differentially expressed genes under these two conditions. The main caveats in analyzing microarray data is that, the numbers of replicates are small and the numbers of genes under study are large also the gene expression may not follow normal distribution. To accommodate all these, we proposed moderated test statistics procedure based on generalized pivotal approach developed to test the hypothesis about the mean of the gene expression levels with the presence of nuisance parameter and also to improve the variance estimation under the circumstances of simultaneous expression of large no. of genes with few no. of replicates. This is formulated based on the assumption that gene expression value follows log normal distribution as an alternative to normal distribution. The improved or shrinkage estimate of variance was obtained by James Stein shrinkage concept, in which the sample variance is shrunk towards its polled estimate resulting in reliable estimates. We also compared our approach with the usual t-test procedure. We applied the above methods to a single channel microarray experiment Golub et al., dataset to find the differentially expressed genes between two types of leukaemia patients (ALL vs. AML). We observed that the number of differentially expressed genes found by generalized p value with shrinkage of variance approach were more as compared to t-test approach for different levels of false discovery rate.

Keywords: Microarray data, generalized p value, t-test, leukemia, false discovery rate.

Paper Name: A Poisson Regression Approach for Association Mapping of Count Phenotypes
Author: Hemant Kulkarni, Saurabh Ghosh

Abstract: While there has been extensive development of statistical methods for association mapping of quantitative endophenotypes governing a clinical end-point trait, development of such methodologies for discrete phenotypes (e.g., symptom counts) remains a challenging area of current research as standard approaches such as ANOVA or linear regression may not be appropriate to analyze count phenotypes. We propose a novel approach based on a Poisson generalized linear
model (PGLM) for association analysis of a count phenotype using data on a random sample from a population as well as on nuclear families. For the population-based design, we consider two tests for association and compare the type-I errors and the powers of these tests with ANOVA, Kruskal-Wallis (KW) and a regression based on a Poisson model. For the family based design, which is restricted to informative trios, we consider a PGLM that models the dependence of the phenotype of the offspring on the alleles transmitted by both parents. We evaluate the performance of four different test procedures based on the alleles transmitted by the two parents. We also compare these tests with the commonly used TBAT procedures (Waldman et.al 1999, Haldar and Ghosh 2015). Our simulations reveal that the Poisson regression does not provide appropriate type-I errors in the presence of overdispersion in the model but the scale adjusted PGLM maintains the proper sizes of the tests. The PGLM model also yields more power compared to ANOVA and KW. For the family based design, inclusion of both parents in the model results in additional power but possibly at the expense of being susceptible to population stratification.

**Paper Name:** Loss to Follow-Up in Prospective Study Designs: Is It A Bad or Good Event

**Author:** Guresh Kumar, Shashi Sharma, S K Sarin

**Abstract:**

Background: In prospective studies, the lost to follow-up and left against medical advice (LAMA) in hospitals is generally considered as bad event like death or non response. In this communication an attempt had been made to model bad events like the death, on responder, dropouts or LAMA cases.

Material and Method: All the cirrhotic patients admitted in the department of hepatology at a tertiary care centre from northern India during the period 2010 to 2015 were enrolled into the study. The stepwise logistic regression analysis was used and a scoring system was developed. The patients were discriminated using threshold value, obtained from the scores developed.

Results: A total of 1728 cases were enrolled, out of which 150(8.7%) were diagnosed as Left LAMA and 391 (22.6%) were deaths. Scoring system was developed based on pneumonia, acute and chronic liver failure, acute kidney failure and GI bleed were important predictors. The discrimination of 150 LAMA cases was classified as deaths in approximately 70% cases. So considering all 150 as deaths can overestimate the results by 2.7% and 30% of cases are misclassified as bad events. In liver diseases, the causality is more and it is estimated that 90% of the dropout/LAMA patients are died. However in acute disease this all cannot be considered as deaths or bad event, in this communication 30% are alive and this can be more in curable disease.
Conclusions: In prospective studies the dropout or LAMA in the hospital should be treated wisely as bad event, depending on the severity and chronicity of the disease; otherwise it will overestimate the bad event rate.

2. Applied Statistics 1

**Paper Name:** Evaluation of Courses under an Undergraduate Academic Program using Quality Indices derived from Data Collected among Student Population

**Author:** Raghavendra Dwivedi, N. N. Pandey

**Abstract:** A set of fourteen Likert type items on a scale of one (strongly disagree) to five (strongly agree) was considered as a structured course evaluation survey (CES) questionnaire to monitor quality of higher education. The data on five Courses taught under B.Com. at P.D. Lions College of Commerce & Economics, Malad (West), Mumbai, were collected from students. For item specific quality assessment, median, first quartile and cumulative percentage with higher grades (4 & 5) were utilized for each item, instead of using inappropriately only mean and standard deviation. The median as four indicates that at least 50% were satisfied whereas as first quartile as four means that at least 75% expressed satisfaction. Using these observed indices, comparative appraisals of quality of courses on each item, global item and also at pooled courses level were made. The appraisal of the pooled analytical results of all courses reveals that less than 50% of students reported to be satisfied on global item. At individual item level also overall satisfaction was not to even that on global item. A mixed picture emerged at individual level analysis under individual courses. As academic planner, aiming to achieve optimal satisfaction among students as 75% and above, each of the subjects needs improvement regarding each of 14 aspects covered under the course evaluation.

**Paper Name:** Time series analysis of human risk due to natural hazards in India and Asia in last four decades.

**Author:** Dr Prabhaker Mishra, C.M.Pandey, Uttam Singh, G.K.Panda Rajanika Kar, Vishal Sharma

**Abstract:**
Introduction: Being the biggest continent in the world, Asia includes around 50 independent countries. Due to specific location, most of its countries are frequently affected by natural hazards and disasters. India is one the country among them. Due to low coping capacity of its people, human casualties are too high. In the present study, a comparative analysis was done for human risk between India and Rest Asian countries.

Objective:
To compare the trend of human casualty due to natural hazards and disasters of India and rest Asia reported in last 40 years.

Materials and Methods: The retrospective study focused on human risk due to natural hazards in India and Asia continent. The data were collected from the website of National Disaster Management Authority, Government of India (EM-DAT, International disaster database). A time series analysis was done to compare the trend. Non parametric methods were used to calculate the significant level between both regions.

Results: Median (inter-quartile range) of the annual human deaths in India and rest Asia was reported 2533 (1753-5813) and 10986 (6376-23829) which was significantly different (p<0.05). Year 2001, 1993 and 1977 are the years where human casualty of India was more than half of the rest Asia.

Conclusions: India contributed on an average 12% human casualty of the Asia continents reported by natural hazards and disasters in last 40 years. There was an annual increasing trend of human casualties of India as well as rest Asia during earlier years but showing decreasing trend in the recent years which is a good news for policy makers and program managers, involved in disaster risk reduction.

Paper Name: Missingness versus multiple imputation – which is the better of the two evils?


Abstract:
Background: In large epidemiological surveys, missing data are a common problem that may threaten both internal and external validity of study inference. Complete-case (CC) analyses that ignore missing data have several problems: being inconsistent, having lower precision, selection bias, and non-nested samples across analyses that prevent comparisons of coefficients across analyses. Multiple imputation (MI), in contrast, can construct a set of complete datasets that include data from all participants and may be used beyond a single analysis. However, MI is not widely used in epidemiological surveys.
Objectives: In the context of a large multi-site survey of South Asian adults, we sought to construct a centralized completed database using multiple imputation, and to compare results from the MI data and CC analyses using dyslipidaemia as a case study.

Methods: Data were from baseline survey of Center for Cardiometabolic Risk Reduction in South Asia (CARRS) cohort during 2010-11. Using multi-stage cluster random sampling, the survey recruited representative sample of 16287 non-pregnant adults aged ≥20 years of urban Chennai and Delhi, India, and Karachi, Pakistan. Multiple Imputation using chained equation (MICE) was used for imputing all missing values across all variables to create ten completed datasets. All covariates and the outcome were included in the imputation model. Imputed values of missing continuous variables were modelled using linear regression and predictive mean matching, and imputed values of ordinal variables were modelled using ordinal logistic regression. Model convergence was checked, and diagnostics were performed on the imputed dataset. In both the CC and MI analysis, we estimated the age and sex standardized prevalence. Multivariate logistic regression was conducted to determine which factors were independently associated with dyslipidemia.

Results: Item missingness ranged from 1% for socio-demographic data to 23% for body weight with 15% in lipids data. The CC analysis included 13707 individuals (84%), while the imputed data utilized data from all 16,287 participants. Univariate distributions of the imputed variables were consistent with those of the observed data. The difference in crude prevalence estimates & standardized prevalence in MI & CC analysis for prevalence estimates by sex and city ranged from (-1.8% to 0.4%) and (-0.9%, 0.7%), respectively. The MI prevalence estimates had smaller standard errors and narrower confidence interval as compared to CC analysis. The difference in Odds ratio estimate in MI and CC analysis ranged from (-0.31, 0.17).

Conclusions: Our work illustrates that a centralized MI database are useful for doing specific analysis. Our analysis of MI data confirmed that the power of MI lies in obtaining smaller standard errors and narrower confidence interval. Also, MI improves the generalizability of the results to the population sampled by allowing inclusion of all individuals in analyses.

**Paper Name**: Confidence Interval to assess “change” or measured on continuous scale when the distribution is skewed

**Author**: Visalakshi Jeyaseelan, Sebastian George and Lakshmanan Jeyaseelan

**Abstract:**

Background: When there are pre and post measurements on same subjects and the outcome of interest is change or to compare the reliability across two methods, then it is required to present mean change and the 95% Confidence Interval (CI) for that
change. However, when the distribution of the ‘change’ is skewed, then it is not possible to calculate CI using normal approximation. This study was to demonstrate an appropriate method in such situations.

Methods: Hypothetical data was considered. Difference of two methods was obtained that included positive and negative values and 95% CI using normal approximation with log transformation, Hodges Lehmann CI, shifting the origin with log transformation and the Bootstrap CI were obtained.

Results: The mean (sd) of the outcome was 96.9 (465.6). The 95% CI using the normal approximation with log transformation was obtained as (245.8, 307.5) with only 194 observations while Bootstrap CI was calculated as (54.1, 139.4) using all observations.

Conclusion: When the outcome of interest is to compare the ‘change’ which is skewed, then we discourage the log transformed normal approximation method or adding constant and taking log transformation method to calculate CI and encourage researchers to use Bootstrap CIs.

**Paper Name:** Factors affecting unmet need of family planning: An analysis of Uttar Pradesh

**Author:** Neha Mishra, Sarvesh Kumar

**Abstract:** In developing countries like India, many married women who say that they would prefer to avoid a pregnancy, are at risk of conception because of the obstacles they face in practicing family planning successfully. These women are said to have unmet need for family planning, a concept that has been identified in recent years to define one of the key population in need of improved family planning services. The measure of unmet need would indicate the proportion of women who fall into the category of need but would also contribute to a greater understanding of its underlying causes- the barriers women face in successfully regulating their fertility. The Study has utilized the secondary data collected under the two rounds of National Family Health Survey conducted in the year 1998-99 and 2005-06 (NFHS-2 & NFHS-3) in Uttar Pradesh. It has tried to define the study variable “unmet need”
under the multinomial regression and its possible determinants. In a population like Uttar Pradesh where there is very high level of infant mortality and also the son preferences, an explicit attempt has been made to examine the effect of sex composition of children on unmet need of family planning. The multivariate analysis for use of contraception using the thirteen selected variables in order of their decreasing importance are urban residency, other caste, other religion, literacy, more than 2 child loss, working women, husbands working in different service category and high standard of living index. From the application of multiple regression models, the important predictor variables of contraceptive use are current age of women, type of place of residence, parity, standard of living index, husband’s education, religion, women’s education, working status of women, women’s occupation, husband’s occupation, caste, child loss and number of living children in order of their decreasing dominance. Out of total selected 13 independent covariates, 12 are found to be significantly.

**Paper name:** Truncated distribution and its modelling using Poisson Regression

**Author:** Vinotha. P, Prem K Mony, Krishnamoorthy Jayanna

**Abstract:**

**Background**

Truncated distributions arise in cases where count data is limited to values which lie above or below a given threshold or within a specified range. Poisson regression is the fundamental model used for analyzing count data which follows poisson distribution. However, using Poisson regression has been shown to have difficulties with zero-truncated data (non-zero counts). Zero truncated poisson model does not perform well in cases of overdispersion. The objective of this study is to demonstrate Zero Truncated Poisson Regression technique on count data in a study to compare documentation of labour room blood pressure monitoring between community health centres and sub-district hospitals of Karnataka state, India.

**Methods**

The data were collected from labour room documentation to assess the feasibility of documentation. Primary outcome considered the number of times (count) blood pressure was recorded from the day of admission to two hours after delivery. Type of facility, type of delivery and length of stay were used as predictors. All statistical analysis was carried out using stata (version 12).

**Results**

There was no occasion of blood pressure not being recorded and hence the data qualified for a zero truncated model. When the poisson and zero truncated poisson
regression results were compared, we observed that the zero truncated model had a better estimation which assessed by Akaike information criterion (AIC).

Conclusion

This study demonstrates the application of zero-truncated Poisson regression in the analysis of non-zero count data. Zero truncated poisson regression is a better approach for non-zero count data.

Keywords: Counts, poisson regression, zero-truncated Poisson regression, maternal care.

3. Health Statistics 2

**Paper Name:** Awareness of health problems during adolescence in girls of Uttar Pradesh
**Author:** Priyanka Sahu, Anjali Yadav, Chandra Pandey

**Abstract:**

Background: Adolescent girls in India while passing through the stage of physical and mental development encounter various health problems. This vulnerable group needs attention of health research groups. Compared to the developed countries, very few data on adolescent health in India is available. Adolescent’s health research is the need of the hour for our country. This will be helpful to stakeholders for strengthening the adolescent health services to cater the related unmet needs.

Objective: The objective of this study is to examine the level of awareness among adolescent girls about physical changes, problems related to menarche, reproductive tract infections, HIV/AIDS and prospective marriage and motherhood.

Methodology: The data for this analysis have been taken from the study entitled Concurrent assessment of health and family welfare programs and technical support to districts of Uttar Pradesh conducted during 2005-2006. Apart from other characteristics, the data on adolescent girls was also collected in the study. The present analysis is based on data from three districts i.e., Rae Bareli, Sultanpur and Faizabad. To assess the level of awareness, adolescent girls were divided into two broad groups assuming that the magnitude of problems and awareness will be different for younger and older adolescents. Further the awareness has been assessed for following five important dimensions of adolescent health (a) physical changes (b) menarche (c) reproductive tract infections (d) HIV/AIDS (e) family formation and motherhood.
Results: About 42.5% girls were aware about onset of menarche and 24.3% about breast development during adolescence. Mainly they receive this information from mothers (44%) and friends (19%). Approximately 50% girls were aware about menstruation, 11% about RTIs and 26% about HIV/AIDS. Main source of information were friends and relatives. About 7% adolescent girls reported to marry below 18 years of age, whereas 40% girls preferred above 18 years. More than half (53%) girls preferred to be married at the age of 18 years which was the legal age for marriage at the time of study. About 46.4% adolescent girls desired first birth during 19-21 years of age. Approximately 47% of the adolescent girls wanted an average birth spacing of 3 years. More than 60% adolescent girls preferred 2 children in their lifetime. The level of awareness about menstruation was high (78%) in older adolescent girls as compared to younger adolescent (26.8%). The analysis concludes that awareness and understanding of health problems during adolescence is associated with age. Major sources of information for any health problem were mothers, relatives and friends. Knowledge and awareness of the issues discussed in this analysis is desired to be universal but knowledge and awareness of these issues among adolescents are always below 50% and in some cases as low as 7% and should be taken up as an integral part of health programmes.

**Paper Name:** A mathematical evaluation of threshold level of 25 (OH) D3 (Vitamin D) in Indians for optimum Bone Health  
**Author:** Ajit Mukherjee  
**Abstract:**  
Introduction and rationale:  
Parathyroid Hormone (PTH) is important for regulation of calcium and Vitamin D in human body and plays a pivotal role in Bone Remodelling. Significant negative correlations between Vitamin D and InPTH were observed both in males ($r = -0.335$, $p<0.0001$) and females ($r = -0.409$, $p<0.0001$) while analyzing data of an ICMR task force study on osteoporosis. It was further observed that sufficient intake of Vitamin D might ensure rise in Bone Mineral Density (BMD) even at low levels of dietary Ca. Thus, it was Important to study the intricate relationship between dietary Ca, Vitamin D and PTH to understand the bone mineral mechanism.  
Objective:  
To determine a threshold level of Vitamin D below which PTH will start rising.  

**Methods:**  
Data were extracted from the larger data set of an ICMR multicentre task force study on osteoporosis carried out during 2001-2006. Expressing PTH as a mathematical function of vitamin D, Loess method of curve fitting was used. The non-parametric Epanechnikov kernel was employed for smoothing the curve in each of the four groups of dietary Ca intake and the point of rise of PTH was assessed.
Results and discussion:
From the four curves that were obtained in each of the four dietary Ca groups, it was quite evident that PTH started rising as soon as Vitamin D levels fell below 20 ng/ml in three of the four dietary Ca groups. However, in the first group with dietary Ca \leq 400 mg/day, the rise in PTH below 20 ng/ml was not so apparent indicating that such low level of dietary Ca may not augur well for good bone health.

Conclusion:
Considering that PTH levels started rising as vitamin D fell below 20 ng/ml, a cut-off point of 20 ng/ml for vitamin D can be safely adopted for ensuring good bone health.

Paper Name: Assessment of Comprehensive Knowledge regarding STDs, HIV/AIDS and Family Welfare Methods among School Adolescents of Varanasi, India
Author: Ujjaval Srivastava, Kaushalendra Kumar Singh

Abstract: A major share of young teenagers is still far from access from education on sexual and reproductive health matters. This cross-sectional descriptive study was conducted within the school going adolescents of Varanasi, India, to assess the knowledge and perceptions regarding family welfare methods, sexually transmitted diseases (STDs) and HIV/AIDS. Here we collected qualitative data showing low familiarity concerning family planning methods, HIV/AIDS transmission and prevention. The acquaintance with STDs is found to be extremely poor and some common misconceptions concerning the transmission of HIV/AIDS are still prevalent among school's adolescents. We have also tested for the association of knowledge with the type of school, medium of school, coeducation and age-group of respondents using bivariate analysis ($\chi^2$-test). It is found that female school adolescents are more ignorant than that of male counterpart on all aspects of the variable under study. There is a need for sexual and health education program to be implemented among school adolescents.

Paper Name: Measurement of inequalities in nutrition and its determinants using Logistic Regression and Concentration Index
Author: Thirupathi Reddy Mokalla, Dr.VishnuVardhana Rao Mendu
**Abstract:**

Objectives: The objective of this analysis is to evaluate the inequality in nutritional status i.e. Stunting, Underweight and Wasting and their determinants in India.

Methodology: This study was performed using the data of National Family Health Survey (NFHS-3, conducted during 2005-2006) and included 41306 children aged 0-59 month. The equality status was assessed by Concentration Index and the relationship of various factors on nutritional deficiencies was assessed by Logistic Regression. These factors included child age in months, mother’s nutritional status, and size of child at birth, wealth index, type of cast, birth order and mother’s nutritional status.

Results: The logistic regression analysis reveals that Stunted children whose age is 13-24 months, OR=4.71, C.I (4.708, 4.280) is the dominant factor affecting chronic malnutrition when compared to children aged 0-6 months, that is followed by mothers educational status with OR=2.2, C.I (2.01, 2.43) when compared to literates, wealth index with OR=2.3, C.I (2.096, 2.518), when compared to Richest, mothers nutritional status with OR=1.453, C.I (1.332, 1.586), when compared to BMI>25.0, size at birth with OR=1.47, C.I (1.074, 1.195) when compared to large size at birth and birth order value with OR=1.27, C.I (1.192, 1.358) when compared to birth order one.

The same results were observed Underweight except children whose age is 25-59 months, OR=2.12, C.I (1.956, 2.298) when compared to children age 0-6(reference) months, mothers educational status with OR=2.42,C.I (2.199, 2.664), wealth index with OR=2.531,C.I (2.304, 2.78), Mothers nutritional status with OR=2.421, C.I (2.199, 2.664), size at birth with OR=1.80, C.I (1.68, 1.92), and birth order value with OR=1.10,C.I (1.04, 1.171).

The same trend was observed in Wasting except children whose age is 0-6 months, OR=2.168, C.I (2.00, 2.35) when compared to children age 24-59 (reference) months, mothers educational status with OR=1.235, C.I (1.097, 1.390), wealth index with OR=1.444, C.I (1.29, 1.61), mother's nutritional status with OR=2.34, C.I (2.07, 2.64) and size at birth with OR=1.65, C.I (1.52, 1.78)

The concentration indexes for stunting, underweight and wasting were (-0.03), (-0.04) and (-0.03) respectively, it indicates that all the three nutrition deficiencies i.e. stunted underweight and Wasted are high in poor children.

Conclusion: The malnutrition is highly dependent on economic status coupled with the factors viz mothers education, mother nutritional status and size at birth.

**Paper Name:** Risk Factors for Pesticide Exposure during Pregnancy  
**Author:** Rakesh Saroj

**Abstract:**
Background: The probable health effects associated with exposure to pesticides during pregnancy are significant public health concern. The Generalized Linear Model was performed to evaluate the impact of pesticide exposure on the risk of preterm birth. All models were fitted in the R software. Methods: The Generalized Linear Model included pesticides exposure as well as other covariates known to affect preterm birth. A survey of a total of 90 women was conducted in SN Medical College, Agra, UP following delivery from January 2015 to February 2016. The dependent variable and explanatory variables used in this study. The Generalized Linear Model graph shows also importance for the study. Result: This study included 90 females (40 subjects of preterm birth and 50 subjects of full-term birth) with relatively homogenous group, having similarity in characteristics such as age, height, weight, BMI, drinking water supply and area of residence. Placental organ chlorine pesticide levels in preterm and full-term cases. The levels of all the pesticides were found higher in preterm cases but the difference were not found statistically significant, only α-HCH found significantly (p > 0.017) higher in preterm cases. The results show other covariate like maternal age (p>0.0051), Chronic Disease, (p>0.017), Baby Weight (p>0.008) and alpha HCH (>0.01721) were highly significant. A Pesticide Exposure was about 1.7 times more likely to have preterm bay (OR=1.67, 95% CI: 1.18, 2.34) compared with a women that was not Pesticide Exposure. Odds ratio were also calculated to estimate risk it is evident that Age (OR= 1.03, 95% CI: 1.01, 1.05), Chronic Disease (OR= 1.4. 95% CI: 1.07, 1.94), Baby Weight (OR= 1.0, 95% CI: 1.1, 1.2) and alpha HCH (OR=0.99, 95% CI: 0.99, 1.00) highly susceptible for preterm birth. Conclusions: In this study we use GLM approach through R Software for evaluates the impact of pesticide exposure on the risk of preterm birth in this population. This study has revealed that Age, Chronic Disease, Baby Weight and AlphaHCH were important risk factors of pregnant women.

Paper Name: Linkages between Consanguinity and Survival of Marriages in Southern India
Author: Himanshu Chaursaisa

Abstract: The Indian subcontinent-a place that is known for plentiful differing qualities, reflected in vast congregation of castes and tribes, excessively depicts changed and confined marriage traditions and norms. Marriage in Indian society is a religious duty. Consanguineous marriage is common, where individuals prefer to marry within their clan. Consanguineous marriages have a higher incidence of divorce, separation, and remarriage than unrelated marriages in India. Divorce has important effects on family and community. In India where consanguineous marriage is common, there is a link between consanguinity and
divorce which is highly important. To the best of our insight, there is no study concerning the linkages between consanguinity divorce and comparison of survival investigation of marriages between consanguineous and unrelated marriages. In this way, the present case-control study was done. The results indicate that consanguinity has some protective role(s) against divorce and also survival of marriages increased among consanguineous marriages. Taken together it might be concluded that consanguinity has social advantages. With the passage of time a general tendency for decrease in consanguinity rate have been found. Several factors like educational status; urbanization and industrialization; deviation from the traditional way of mate selection etc. plays a role in minimizing the trend. On the other hand, consanguineous marriages are associated with increased risk of recessive traits and also it might be associated with many multifactorial diseases. However, activities for reduction of consanguinity which is culturally favoured in population, without attention to its social reflections, are not recommended.

Session (3.50-5.05)

1. Sampling

Paper Name: Estimation of population mean in two stage sampling using successive sampling technique under non response

Author: Gautam Vishwakarma

Abstract: A survey could be an ad hoc or a repetitive. An ad-hoc survey is one, which is conducted without any intension of or provision for repeating it, whereas a repetitive survey is one in which data are collected periodically from the same, partially replaced or freshly selected sample units. If we want to study about the current occasion, the survey can be an ad-hoc one or one-point survey. But when changes occur on some characteristics over time are of interest, it is necessary to carry out the survey repetitively. The main aim of this paper is to discuss the occurrence of non-response in the population on the current occasion while dealing with sampling over two occasions under repeated surveys in two-stage sampling and to suggest some sampling strategies which could be adopted under such situations. While replacing a part of the sample of the current occasion through matching process, we have considered estimation procedures under the assumption of non-response in fresh, in matched and in both the samples. Properties of these strategies have been observed. In support of the theoretical considerations, an empirical study has been presented at the end.
**Abstract:** In recent years both in academic and industry environment the phrase “Big Data Analytics” is becoming popular and also becoming familiar to the new generation of Statistician and other allied supporting professionals (like Mathematician, Computer Science Scientists etc.). The newly developed subject Big Data Analytics can be defined as a combination of many subjects like computer science, scientific software, Mathematical algorithms, Statistical optimization/techniques along with functional area of interest (Economic, Marketing, Biology, Health care, Demography etc.).

Big Data Analytical research is already started in big schools both in India and abroad. In graduate schools, are also offering the Masters program in this area. But there were no clear cut methodology incorporated yet. Here, we are in an attempt to familiarise some of the Statistical techniques are commonly useful in this area of research. I will demonstrate these techniques with some of the examples taken Health Survey data as a functional area of interest.

**Abstract:** In this paper we have proposed an exponential ratio type estimator of the finite population mean in presence of measurement error and non-response error. Expression of Mean Square Error has been calculated up to first order of approximation. Empirical study based on five data sets has been carried out. Theoretically and empirically we have found that proposed estimator is more efficient than the existing estimators in presence of measurement error and non-response error. The properties of estimators are generally analyzed under the supposition that observations have been recorded without any error. But this is not in actual practice and the data may contain error due to many reasons. These errors make the result invalid. If these errors are small, it can be neglected. The term non-response in sample surveys is the failure to get information from some units of the population due to various reasons like unavailability of respondents, lack of information and refusals etc. Non response arises when the respondents chosen for the sample are not able to participate in the survey. It is clear from our result that our proposed estimator has largest PRE (percentage relative efficiency) i.e. is the
most efficient estimator among all other estimators that we considered in this paper. We have also computed PRE for estimators in presence of measurement error and in absence of measurement error. From our findings we conclude that estimators show unexpected increase in efficiency when measurement error is not considered. The PRE’s (in case of without measurement error) becomes approximately double than the PRE’s (in case of with measurement error).

There is huge amount of difference between efficiency of estimator in presence of measurement error and when it is not present. It is also clear that how the percentage relative efficiency changes as sub-sample size i.e. value of \( h \) changes. We can also observe that efficiency of proposed estimator is much affected by size of sub-sample when measurement error is not considered and when measurement error is considered efficiency is not as much affected by sub sample size as in the opposite case.

From our empirical study we draw result that sometimes measurement errors are not negligible as they affect efficiency of estimators at high rate. Its ignorance may cause undesirable or unexpected consequences. As in the case of estimator, it performs better than when measurement error is considered and when measurement error is not considered result is totally reversed, and performs better than. Our proposed estimator is much more efficient among all the estimators that we have considered here. Hence it could be used for future use. And also it is advisable to perform estimation in presence of non-response and measurement error.

**Paper Name:** Estimating Population Mean In Systematic Sampling Using Two Auxiliary Variables

**Author:** Madhulika Mishra

**Abstract:** It is widely known in the context of literature of survey sampling that the efficiency of the estimators of the population parameters of the variable of interest can be increased by the use of auxiliary information related to auxiliary variable \( x \) which is highly correlated with the variable of interest \( y \).

Hansel(1942) and Griffith(1945-1946) found systematic sampling to be more efficient and convenient in sampling certain natural populations like forest areas for estimating the volume of the timber and areas for estimating the volume of the timber and areas under different types of cover(Osborne 1942). Systematic sampling is generally more precise than simple random sampling and even more precise than stratified random sampling under certain specific conditions. Also it provides estimators which are more efficient than simple random sampling or stratified random sampling for certain types of populations see [Cochran (1946), Gautschi(1957), Hajek(1959)]. Various authors like Kushwaha and Singh (1989), Banarasi et al. (1993), Singh and Singh (1998), Singh et al. (2011), Singh and Solanki (2012), Singh and Jatwa (2012, Singh et al. (2012a), Singh et al. (2012b),
Chaudhary et al. (2012), Verma et al. (2012), Verma and Singh (2014), Singh and Singh (2015) and Khan and Singh (2015) have discussed the problem of estimation of population mean using information on auxiliary variables in the context of systematic sampling. In this paper, we have proposed an estimator for the estimation of population mean using information on two auxiliary variables under the framework of systematic sampling. Through this manuscript, the problem of estimating the population mean in stratified sampling using information on two auxiliary variables has been addressed. The expressions for the mean square error of the proposed estimator have been calculated up to the first order of approximation. It has been shown that the proposed estimator is more efficient than other existing estimators for which an empirical study through a real secondary data has also been carried out.

Paper Name: An Improved Generalized Class of Estimators for Population Variance Using Auxiliary Variables
Author: Nitesh Adichwal
Abstract: In this paper we have proposed an improved generalized class of estimator for estimating population variance based on simple random sampling without replacement. The expression of bias and mean square error of the proposed estimator are obtained up to the first degree of approximation. We have derived the conditions for the parameters under which the proposed estimator are more efficient than the usual estimator and other existing estimators. An empirical study is also carried out using two real population data sets in the support of theoretical results.

2. Applied Statistics 2

Paper Name: Comparison of Imputation Techniques in Missing Data Analysis: A Simulation Study
Author: Seena K, Dr.Thennarasu K, Dr.Mariyamma Philip
Abstract:
PURPOSE: The problem of missing data is challenging in Statistical Data Analysis. It offers infinite scope for finding out easy and efficient methods for handling the situation.
OBJECTIVE: To examine the performance of imputation methods, Quantile Regression imputation(QR), Multiple Imputation(MI), Regression Imputation(RI), Stochastic Regression Imputation(SRI) and Mean Imputation (MeanI) comparison with Full data(FD) and List Wise deletion method(LW).
METHODOLOGY: This is a simulation study in which performance of the methods were evaluated based on different sample sizes 30, 50, 100, 200 and 500, on a dependent variable (DV) and two independent variables (IVs). Two data generating models were used namely the simple linear location-shift model and the heteroscedastic model. One of the IVs follows uniform distribution and the other follows chi-square distribution independently. In both models error was generated from chi-square distribution. The missing values were created for one of the IVs under MAR mechanism and for other under MCAR mechanism so as to generate 50% proportion of missing. Under the location shift model, Quantile Regression (tau=0.1 and 0.5) as well as Linear Regression model were fitted for the Full data, List wise Deletion and five imputation methods. Quantile Regression (tau=0.1 and 0.5) was performed for the second model and linear regression was not done due to the heteroscedasticity. In each setting, 1000 simulations were done and the average estimates of the regression coefficients, their errors, variance ratio (VR), bias and mean square errors were compared for all the methods of imputation.

RESULTS: For the data generated under the homoscedastic model, the average estimate of the intercept for the FD was 1. The MI method gave better estimate with an average value of 1.006 (VR=1.584) for the sample sizes 30 to 500 with least bias for the intercept -0.005 (SE=0.136). Under the homoscedastic model, the average estimate of the intercept of Quantile Regression (tau=0.1) for the full data was 0.206 (SE=0.124). The better estimate was given by QR method with an average estimate of 0.212 (VR=3.249) with least bias of -0.006 (SE=0.206). Under the heteroscedastic model (tau=0.1) the average estimate of intercept of FD was 0.179 (SE=0.321). QR and MI methods gave better estimates with average values 0.124 (VR=2.525) and 0.107 (VR=2.174) in which MI has comparatively less VR, least bias given by QR with an estimate of 0.056 (SE=0.446).

CONCLUSION: Imputation methods were found to be superior to LW method in all the situations of analysis. For the data generated with homoscedastic model, the average estimate of the linear regression coefficients of MI method showed better results compared to other imputation methods. For the quantile regression analysis, QR method gave better estimates for the same model. Under the heteroscedastic model the methods QR and MI produced similar results with least values of bias and variance ratio.

Paper Name: Job Satisfaction and its Related Factors among Nursing Staff in a University Hospital of Varanasi
Author: Subhi Srivastava, Alok Kumar
Abstract:
Background: Job satisfaction is generally regarded as an employee’s attitude toward the job and job situation. Job dissatisfaction leads to absenteeism, labour turnover and negative publicity of the organization. Unsatisfied worker can prove to be a liability to any organization and on the other hand happy and satisfied worker is
always a productive worker. Nurses contribute very important position in health care sector. A nurse is responsible—along with other health care professionals—for the treatment, safety, and recovery of acutely or chronically ill or injured people, health maintenance of the healthy, and treatment of life- threatening emergencies in a wide range of health care settings. Nurses may also be involved in medical and nursing research and perform a wide range of non-clinical functions necessary to deliver health care. Thus, studying job satisfaction among nurses is important because it is necessary to distinguish between positive and negative aspects of job satisfaction.

Objective: The study aimed to explore the views and experience regarding different components of nurse’s working lives and analyse various factors influencing job satisfaction among nursing staff in a university hospital of Varanasi.

Methods: A cross-sectional survey design utilizing questionnaires will be selected to fulfil the research objectives. A total of 163 nurses working in different medical and surgical departments in the Sir Sundar Lal Hospital of Banaras Hindu University in Varanasi completed the questionnaires.

Results: About 60 per cent of the nurses said that nursing job were the first choice of their career. Less than half of nurses (41.2 per cent) were satisfied or very satisfied with their jobs and about 60 per cent felt moderate to extreme occupational stress. Half of the sample reported a high level of organizational commitment (49.1 per cent) and only one-tenth of the respondents accounted professional commitment (10.9 per cent). Only 14.5 per cent reported role conflict often or very often. About two-third of nurses (64.1 per cent) were satisfied or very satisfied with physical conditions in which they work. Nurses with a diploma or associate degree reported greater professional commitment and a lower level of role conflict than those with a bachelor degree (p<0.05). Married nurses reported more organizational commitment compared to unmarried and divorced. Middle year age-group (25-35) of nurses significantly reported high level of organizational commitment and often to very often role conflict as compared to young and old age-group of nurses.

**Paper Name:** Monte Carlo analysis of multilevel binary logit model to improve the prediction  
**Author:** Bhaskar Thakur, Vishnubhatala Sreenivas, S.N. Dwivedi, Mona Pathak  
**Abstract:**  
Background: Many epidemiologist, public health researcher and social scientist aim to quantify the contextual phenomena at county level when the outcome belonging to the binomial family but not realize the general problem of reliability of the estimates when there is small number of counties. They apparently use the conventional method of either Laplace approximation or quasi-likelihood estimation. The performance of adaptive quadrature is better in the computer based Markov Chain
Monte Carlo (MCMC) method, which is believed to obtain more accuracy on random effect variances. The aim of this work is to compare the performance of the variance component model based on the different likelihood estimations as well as MCMC method.

Method: The risk of death in post-neonatal period was estimated through the different likelihood estimation procedures and MCMC method among the representative sample of 13612 women interviewed in Bihar state was been drown from District Level Household Survey3 in India. Infant and district were considered as first and second level respectively in hierarchical data structure. The rank correlation of observed and model predicted post-neonatal mortality and Area under the ROC curve was compared based on different estimation procedures to assess the performance.

Result: Although the district level variance is very small but estimated differently when maximum quasi-likelihood of first order (0.061), penalized quasi-likelihood of first order (0.060), penalized quasi-likelihood of second order (0.065) were used. Lowest value of district level variance (0.049) was estimated when Markov Chain Monte Carlo (MCMC) was used. Rank correlation is found to be higher in case of MCMC procedure as compared to other classical estimation procedures.

Discussion: Performance indicators such as rank correlation and discriminating ability as assessed by the area under the ROC curve clearly demonstrated the utility of MCMC method in multilevel model in hierarchical data structure.

Paper Name: Analyzing correlated outcomes with outliers
Author: Nivya George, Tinku Thomas
Abstract:
Introduction: We often deals with non-normal data and outliers in many fields. The conventional approaches are using normalizing transformations and rank based transformations. Even after the transformations, the data may not follow normal distribution and the outlying observations may alter the results of overall analysis. We need a method which gives less weight to outlying observations to get the correct results. We explore the use of SAS procedure GLIMMIX which is a linear mixed model approach to estimate the parameters in such data.
Method: A small sample data with repeated reading and outliers was analyzed using Wilcoxon Signed rank test, paired sample t test on the log transformed values were done to compare the iron absorption in a cross over randomized trial. The estimates from linear mixed model (using GLIMMIX) were examined and compared.
Results: The non-parametric Wilcoxon Signed rank test provides a simplistic approach to analyze data with outliers without considering the study design characteristics. However it is not possible to compute an effect size from such
analysis. The paired t test of log transformed values though provides an effect size, could probably give an over or underestimate due to ignoring the outlier. Utilizing certain options in PROC GLIMMIX provides a direct and more precise estimate of the effect size while considering the presence of outliers and complex study design. There findings are described using an example.

Conclusion: GLIMMIX procedure can be used to handle the outliers in data arising from complex design such as correlated outcomes.

Keywords. GLIMMIX procedure; outlier

**Paper Name:** Evaluation of the utility of FlexMix and CAMAN packages in Mixture Distributions

**Author:** Bijesh Yadav, Sebastian George and Lakshmanan Jeyaseelan

**Abstract:**

A mixture model is a probabilistic model for representing the presence of subpopulations within an overall population, without requiring that an observed data set should identify the sub-population to which an individual observation belongs. The most widely used software to analyze such data is R and SAS of which R is preferred as it is freely downloadable. Also, there are extensive codes available to analyze mixture models using R. However, R has different packages to analyze such Mixture models which are ‘Mixtools’, ‘FlexMix’ and ‘CAMAN’. However, it is important to know under what situations or conditions these packages provide similar or different findings.

Objective: The main aim of this study was to compare the findings from a normal mixture generated data using FlexMix and CAMAN packages by assessing their BIC, Bias, mean square error values for simulated and real time data.

Material and methods:

Consider finite mixture models with K components of form

\[ h(y|x, \psi) = \sum_{k=1}^{K} \pi_k f(y|x, \theta_k) \]

\[ \pi_k \geq 0, \quad \sum_{k=1}^{K} \pi_k = 1 \]

Where y is a (possibly multivariate) dependent variable with conditional density \( h \) (h=Normal), x is a vector of independent variables, \( \pi_k \) is the prior probability of component k, \( \theta_k \) is the component specific parameter vector for the density function...
f, and \( \psi = (\pi_1, \ldots, \pi_k, \theta'_1, \ldots, \theta'_k) \) is the vector of all parameters. The estimation of these parameters is based on EM algorithm.

The ‘mixcov’ function of CAMAN package and ‘flexmix’ function of FlexMix is an interface to R’s generalized linear modelling facilities.

Data simulation:

Mixture of linear Regression data: Covariate was generated from the specified sequence and Outcome was generated from the two or three normal distributions with the given means and standard deviations using the rnorm () function with varying sample size.

Assessment of package was based on BIC, Bias and Mean square errors and coverage probability. Lower BIC values, bias and mean square errors were considered to be better models where as coverage probabilities nearing 95% were considered as better model.

Results:

The “b” values were almost similar in both the packages with different samples sizes. BIC values increased with increasing sample sizes as known. The coverage probabilities were slightly higher in FlexMix Package as compared to the CAMAN Package with different sample sizes. However, model accuracies like MSE and bias were lower in CAMAN package as compared to the FlexMix. BIC’s were also consistently lower in CAMAN package irrespective of sample sizes as compared to FlexMix Package.

Conclusion: CAMAN package has been consistently better when applied.

---

**Paper Name:** Misuse of Harmonic Regression method for Time series data

**Authors:** Kavitha R, Sebastian George and Jeyaseelan L

**Abstract:**

Time series data exhibits specific qualities like trend, autocorrelation and seasonality. Analyses of such datasets handled in two ways. Either using time series based methods like ARIMA (Auto Regressive Integrated Moving Average) or using regression based methods like Harmonic Regression (HR), adjusting for the qualities like trend and seasonality.

**Objective:**
To demonstrate the misuse of the Harmonic regression method for time series data by simulating data.

Methods:

In Time series (ARIMA) based simulation, we have generated data with 240 observations each for AR values from 0.05 to 0.95 by 0.05, which had 19 datasets in total. And to simulate data based on regression method, we simulated based on HR model. The values for coefficients of Sine and Cosine were taken as values of 0.5 to 10 incrementing by 0.5. We have generated data with 240 observations in each loop. Then, we evaluated HR and ARIMA models fit to the simulated data.

Results:

In ARIMA based simulation, the MSE (Mean Square Error) and MAD (Mean Absolute Deviation) values for ARIMA model is much lower than HR as expected, because simulation is based SARIMA model. In HR bases simulation, the MSE and MAD values for HR model were a bit better than ARIMA as expected, because simulation is based HR model. But, this is true only if we simulate HR data only with seasonality. When we add a term for trend in the HR simulation, and then fit HR and ARIMA models for the data, ARIMA had lesser MSE and MAD compared with HR.

Conclusion: Though Harmonic regression method is commonly used for analysing time series data, the appropriateness is true for data with only seasonality but not for data with trend.

3. Health Statistics 3

**Paper Name:** Application of Multimodal count regression to predict the number of Somatic symptoms of pregnant women  
**Author:** Adhin Bhaskar  
**Abstract:**  
BACKGROUND: Women are likely to experience changes in physical and mental health during the period of pregnancy. The symptoms vary from nausea, weakness, tiredness, backache, appetite changes, etc. The current study intends to identify the factors affecting the number of somatic symptoms associated with pregnancy.  
AIM: To identify the factors responsible for the number of somatic symptoms among pregnant women using appropriate count regression model.  
MATERIALS & METHODS: Poisson and Negative binomial regressions are the two obvious approaches used to study the relationship of count response variable with the predictors. This study is focused on comparing the performance of various count
regression models when the dependent variable is both overdispersed and multimodal. This is a community based prospective study of anxiety, depression during pregnancy and its relationship to pregnancy outcomes. The data was collected from 609 pregnant women attending antenatal clinic, Bengaluru. Poisson, Negative Binomial, Hermite and generalized Hermite regression models were fitted to find the relationship between the variables. Haritgans dip test detected the existence of multimodality in the response variable. Moreover, the auxiliary regression based test which was administrated after fitting the Poisson regression indicated the presence of overdispersion in the data. The models were compared using Likelihood ratio test, Root Mean Square Log Error (RMSLE) and Akaike Information Criteria (AIC) along with the estimates and standard errors.

RESULTS: The generalized Hermite regression was chosen as the best to establish the relationship between the response variable, number of somatic symptoms and the predictors based on the values of fit indices and tests. The model identified education, abortion, stress and depression as the factors affecting number of somatic symptoms in pregnant women. The study found out that depressed women had 44% more risk to experience more number of somatic symptoms as compared to non depressed women.

CONCLUSION: The incidence of overdispersion caused the Poisson regression to overestimate the relationship between the variables. The generalized Hermite regression outperformed other conventional methods and standard Hermite regression in modeling multimodal and overdispersed count data.

Paper Name: Determination of Threshold Value/S from A Battery of Psychological Tests for the Diagnosis of Dementia.

Author: Palaniappan Marimuthu, Ravikesh Tripathi, Keshav Kumar, Srikala Bharath

Abstract:

Abstract. Aging is a global phenomenon and the proportion of older persons is increasing across the globe. In absolute terms, the number of elderly has tripled over the last 50 years (United Nations, 2009). Differential changes in the spectrum of cognitive functions have been identified in the process of normal aging, and in clinical conditions such as mild cognitive impairment (MCI) and dementia (Rediess & Caine, 1996). Decline in episodic memory (particularly delayed recall) is usually the earliest cognitive change that occurs prior to the development of AD dementia (Salmon & Bondi, 2009).

The Objective of this study is to develop a battery of neuropsychological tests and standardize it for older adults in India.
The normative sample (N=180), in the age range of 55-64 years meeting the inclusion and criteria and clinically proved MCI/AD were 49 subjects selected for this study. Following test were developed for testing the different cognitive functions of the brain: Word List, Story Memory Test, Stick Construction Test, Tower of Hanoi, Fluency, Go/No-Go, Digit span, Cors block-tapping test, NIMHANS parietal focal sign.

Education was found to be a strong contributor with significant effect on all variables, which accounted for 3%-48% of variance on each task. Except on the Stick Construction, the contribution of education was very high, that is greater than or equal to 50% of the variability explained on the dependent variables. This implies that the role of education is important in explaining the variability on test performance followed by gender and age. The area under the ROC curve (AUC) is used to determine discriminatory ability of the tests, it signifies how well any particular test discriminates between individuals with dementia and controls. The results indicate that Word List delayed recall and story memory delayed recall have good accuracy with AUC of 0.97 and 0.90 respectively. Word List learning trial three, Word List learning total, stick construction delayed recall, Story Memory immediate recall, Word List learning trial two, Fluency for Animals, Fluency for Vegetables, Word List learning trial one, Fluency for Fruits, Word List recognition were found to have AUC between 0.70 to 0.86.

The neuropsychological battery measures attention, working memory, episodic memory, semantic memory, construction and parietal focal signs. Education emerged as a significant determinant, influencing the neuropsychological test performance. The test battery could discriminate well between cognitive impairment and normal older population. Finally, a structural equation model was used for dementia group to understand the effect of age, education and gender on the digit span (reverse) test.

**Paper Name:** Human deaths due to natural hazards in Bangladesh  
**Author:** Vishal Sharma  
**Abstract:**

Background: Natural hazards and disaster is one of the major causes of the human death in Bangladesh. Due to its specific location, it is frequently affected by natural hazards and disasters including flood and cyclone which is one of the most costly disasters in terms of both property damage and human casualties. Despite huge attempt made by disaster managers, reducing human casualty is still a major challenge for them. Present study focused on human deaths due to natural hazards and disasters reported in last 50 years in Bangladesh.
Objective: To discuss on trend of human casualty due to natural hazards in last five decades (1966-2015) in Bangladesh.

Materials and Methods: The retrospective data of human deaths due to natural hazards in Bangladesh were used to analysis. Data were collected from the website of National Disaster Management Authority, Technical Reports, Government of India (EM-DAT, International disaster database). Independent samples median test used to compare median human casualty, among five decades.

Results: During 1966-2015, natural disasters were occurred in 49 years. Minimum human casualties were recorded, 50 in 2013 and maximum around 3 lakhs in 1970. Median (interquartile range) of the human casualty was 666 (231-1536). There was a non-increasing trend of human casualty over the last five decades. Independent samples median test shows that there was significant difference in median human deaths over the decades (p<0.05).

Conclusions: After many preventive measures, overall human casualties are showing decreasing trend in Bangladesh. There is a need to keep on disaster mitigation and risk reduction programs with increasing coping capacity of the people to minimize the human risk at lowest level.

Paper Name: Prediction of anxiety and depression among the sailors using machine learning technology
Author: Arkaprabha Sau

Abstract:
Introduction: Shipping industry is indispensable for economic growth and sustainable development of the society. Sailors are the most important member of that industry. Due to work schedule, life style, job profile, and other factors, they are more vulnerable to suffer from various mental disorders, most commonly anxiety and depression. So, periodic screening, followed by necessary treatment for anxiety and depression, is necessary for their physical and mental health and wellbeing.

Occupation health specialists have to do this job by interviewing the sailors with standard predesigned and pretested schedule like Hamilton anxiety and depression rating scales (HAM A and HAM D), Hospital Anxiety and Depression Scale (HADS) etc. It takes approximately 30 minutes to complete an interview for this purpose only. As an alternative, machine learning technology can be used as a quick and automated screening procedure to identify the at risk sailors for early referral to psychological counselling and treatment centre.

Methodology: An observational descriptive study was carried out among 470 sailors at Haldia Dock Complex under Kolkata Port Trust, Ministry of Shipping, Government of India, between January 2016 to August 2016 after taking necessary permission and ethical clearance. Various socio demographic, occupational, and health related
information were collected. Then status of anxiety and depression was assessed by Hamilton Anxiety and Depression rating scale. Different attributes like age, marital status, job profile, employment status, working hours, duration of service, chronic disease condition, etc. were selected to predict one of the 4 different outcomes i.e. Normal, Only Anxiety, Only Depression, Anxiety and Depression both. Five machine learning classifier i.e. Bayes Net, Logistic, Multi-Layer Perceptron, Decision Table, and Random Forest, were evaluated using WEKA, a data mining tool used for machine learning based prediction. Principal component was selected as an attribute evaluator with Ranker search method.

Result: After evaluating 5 machine learning classifier with 10-fold cross validation test option, Random forest appeared to be the best one for this purpose with true positive rate of 0.881, false positive rate 0.065, precision 0.883, and ROC area 0.964. Among 470 sailors this model correctly classified 414 sailors, with kappa statistics 0.81, mean absolute error 0.876, and root mean squared error 0.2099.

Conclusion: Using this machine learning technology, time consuming, manual, anxiety and depression screening procedure by various rating scales, can be replaced by an automated computer based technique with reasonable amount of accuracy. This will be beneficial both for the sailors and occupational health specialists. This study was carried out only one major port in India. Similar kind of large cross sectional studies involving others ports in India, is necessary to develop a generalize model for future application.

**Paper Name:** Mentoring For Addressing Reproductive And Sexual Health Needs Of Unmarried Adolescents In Chandigarh (Ut), India  
**Author:** Dinesh Kumar  
**Abstract:**
Context: Reproductive and sexual health challenges of adolescents and their growing needs have so far been neglected in India. Objectives: To investigate reproductive and sexual behaviour of adolescent students and to evaluate impact of mentoring on their reproductive and sexual health needs. Methods: A Longitudinal survey among 1819 adolescents in schools of Chandigarh, India, selected by stratified multi-stage random sampling. Mentorship program was introduced in selected schools. They were interviewed to collect information on reproductive and sexual issues. Results: In pre-intervention survey, 26.2% respondents had some intimate friends, 81.5% were exposed to sex related material and 73.6% were aware of sexual abuse. In post-intervention survey, 32.4% adolescents reported intimate friendships, 93.2% had exposure to sex related material. Results of pre-intervention survey indicated awareness of teenage pregnancy among 62.6% respondents and awareness of premarital sex among 43.9% adolescents. In post-intervention survey, 75.0 % adolescents had knowledge about premarital sex and 52.5% were of the
opinion that it was dangerous and should be practiced only after marriage. Overall contraceptive awareness (78.3%) observed in pre-intervention survey was increased in post intervention survey to the levels of 84.7%. Awareness levels of STI’s and their modes of transmissions, preventive measures etc. were also increased reportedly. Felt need of sex education in schools was on rise from 60.8% to 74.2% .

Conclusions and Suggestions: The study found several reproductive and sexual health challenges and suggests the need of introducing “Mentoring” as an effective Adolescent and Youth Friendly Health Initiative (AYFHI) strategy at school levels. Clinical Implications and National Relevance: Mentoring during adolescence may strengthen the Adolescent Health Programs by creating reproductive health awareness among them for Planned Parenthood. It can successfully address Adolescent Reproductive and Sexual Health (ARSH) challenges.

**Paper Name:** Trend and Risk factor analysis of Singleton and Twin Caesarean deliveries.

**Authors:**

1. Tunny Sebastian MPhil, Department of Biostatistics, Christian Medical College & Hospital, Vellore
2. Visalakshi Jeyaseelan PhD, Department of Biostatistics, Christian Medical College & Hospital, Vellore
3. Jeyaseelan Lakshmanan PhD, Department of Biostatistics, Christian Medical College & Hospital, Vellore
4. Sebastian George PhD, Department of Statistics, St. Thomas College, Palai.
5. Reeta Vijayaselvi MD, Department of Obstetrics and Gynecology(Unit 4), Christian Medical College & Hospital, Vellore,
6. Ruby Jose MD, Department of Obstetrics and Gynecology(Unit 4), Christian Medical College & Hospital, Vellore, India

**Abstract:**

Objective: To analyse the increasing and decreasing trend and risk factors of singleton and Twin Caesarean Deliveries (CD) over 11 years.

Methods: A retrospective analysis of CDs between 2000 and 2010, with gestational age between 28 and 42 weeks was done. The overall CD trend as well as the risk factors for singletons and twin pregnancies were analysed. The sequence charts and segmented regression analysis were performed for the trend analysis of the CDs. The risk factor analysis of singleton deliveries was performed using univariate and multivariate logistic regression analysis and of the twin deliveries were performed
using Generalized Estimating Equations analysis. SPSS and STATA softwares were used for the data analysis purpose.

Results: A significant change in trend for CD: a steady rise (10.2%) from 20.9% in 2000 to 31.2% in 2006, followed by a slow rise (2.8%) to 33% from 31.2% between 2006 and 2010, was observed when simple interventions to reduce CD rate was instituted. CDs among twin pregnancies did not show the same trend.

Among singleton deliveries, demographic factors of increased age, height less than 150cms, higher body mass index, poorer levels of education of women and presence of obstetric factors of hypertensive disease, antepartum hemorrhage, abnormal Doppler indices, preterm premature rupture of membrane, decreased amniotic fluid index, clinical and sub clinical oligoamnios, Chorionamnionitis, breech or transverse lie at delivery, Preterm delivery, small and large for gestational age babies and male babies revealed a significantly higher risk for CD. Infertility and Previous CD also showed a higher risk for CD.

Among twin pregnancies, a higher body mass index and non vertex presentation at delivery were significant risk factors. Teenaged women and preterm delivery also had a higher risk of CD among twins.

Conclusion: Segmented regression analysis is a useful tool for finding significant difference in trends in intervention analysis. Simple interventions can be instituted to decrease CD rates. Preterm births constitute a majority of CD; hence measures to decrease preterm birth should be a priority.
model the explanatory variable or covariates are used for the random effects of covariates on distribution of survival time. In this article, we considered the base line distribution of variable as Exponential distribution. Considering this we estimated the frailties parameters.

This article contains the estimation of survival function using linear hazard function and exponential base line distribution. Considering PH model with exponential base line distribution, the maximum likelihood estimator of general linear parametric function of regressors and frailty parameter are obtained by taking very general form of the regression matrix. It is observed that estimator is biased and hence MSE is also obtained.

The result is supported using Kidney infection data.

**Paper Name:** Modelling and Forecasting In Seasonal Epidemic Surveillance Data
**Author:** Manik Awale

**Abstract:** Most of the epidemics are seasonal in nature, e.g. influenza, cold, mumps, chickenpox etc. The seasonality in the disease spread appears in the epidemic surveillance data and therefore one needs to account for this while proposing time series models. It appears that the susceptibility, infectivity and immunity of individuals also changes with the season and hence, the probability of getting infected or otherwise changes with seasons and time. In this paper, I have considered the time changing probability of infection to model the seasonality using integer valued autoregressive model of order one (INAR(1)). An attempt is made to allow the thinning parameter to change with time using trigonometric functions to address the seasonality in disease progression. This model results in better forecasts of the epidemic, as compared to the non-seasonal INAR(1) model. The estimation of the parameters is done using conditional maximum likelihood (CML) and conditional least squares (CLS). The coherent forecasting is carried out for simulated as well as real data sets which are having seasonality. The suggested model addresses the seasonality more logically as compared to the existing models.

**Paper Name:** Evaluation Of Diarrhoea Endemic And Non- Endemic States Of India: A Spatio Temporal Analysis
**Author:** Nilima Nilima, BINU V S, Karthik Shetty

**Abstract:**
Background: Diarrhoea is a major cause of mortality in children under 5-years age. It is the second most alarming public health problem in India which Government has been trying to monitor & control. With an Objective to obtain the hotspot and the Spatio-temporal pattern of risk factors associated with diarrhoea across the endemic and non-endemic regions in India. Methods: Data is obtained for ten (5- endemic and
non-endemic) and eight (4-endemic and non-endemic) states from 5 different regions of India, from District Level Health Survey -3 and 4 respectively. We used Moran’s I and LISA to detect the spatial clustering of diarrhoea cases across the states of India and to test for clustering in the data. Spatial regression is used to identify the factors associated with diarrhoea in endemic and non-endemic states of India. For analysis we used GeoDa and SPSS. Results: The Moran’s scatter plot for DLHS-3 suggests the spatial pattern in the distribution of diarrhoea in India whereas In DLHS-4, no spatial pattern in distribution of diarrhoea was observed. Using the LISA cluster map, in both data, we identified that Maharashtra is a consistent hotspot for diarrhoea cases and also the factors named children suffering from diarrhoea who were given treatment in the last two weeks and use of piped drinking water are found significant over time. Conclusion: This study exhibits that the Special care should be provided to Maharashtra state in order to reduce the number of diarrhoea cases significantly by taking suitable measure towards source of drinking water and proper care need to be taken towards the timely treatment provided to children with diarrhoea.

Paper Name: Analytical study of awareness and behavioural practices for prevention of malaria in the high endemic areas of Assam, India.
Author: Himanshu K. Chaturvedi, Preeti Tiwari and Arvind Pandey

Abstract:
Background: The present analytical study was carried out to demonstrate the contribution of behavioural practices for prevention malaria in the high endemic areas. The survey data on health seeking behaviour has been used and analysed to assess the factors influencing the behavioural practices to prevent malaria.
Methods: A sample of 1989 households’ survey data collected from high endemic areas of Assam was analyzed using multinominal logistic regression model to determine the important predictors associated with behavioural practices.
Results: The average age of respondents was 41.1+ 12.0 years and 71% of them were males. The analysis indicates the strong association of high and medium level of behavioural practices (with reference to low level) with all other variables such as age, education, religion, occupation, type of house and level of awareness. The level of awareness and behavioural practices of malaria was 35% and 30% in high level, and 48% and 22% in medium level, respectively. However, high awareness is not resulted completely too high behavioural practices as expected. Overall, the analysis indicate that medium and high level of behavioural practices related to prevention of malaria was highly associated with awareness, personal background and living standard of people.
Conclusion: There is need to improve the awareness and behavioural practices of people for prevention and control of malaria by proper health education especially in high endemic areas of Assam, India.
**Paper Name:** Burden of Cancer in India in terms of absolute number  
**Author:** Atul Juneja, D. Sahu, A. Sehgal, A. Pandey

**Abstract:** Burden of disease can be considered through various approaches depending on the need of the program. In view of demographic transitions non communicable diseases like cancer have become an important candidate for being studies for disease burden. The present exercise attempts to compute the absolute number of cancer cases in Indian urban population based on the data generated through National Cancer Registry Programme of Indian Council of Medical Research (ICMR). The age specific data on cancer of cancer incidence for the urban registries Bengaluru, Bhopal, Chennai, Delhi, Mumbai, Delhi have been used with the urban population data of the country as available through Registrar General of India. This latest census (2011) data has been smoothened and the category where the age was not stated was distributed proportionately. It was estimated that the load of cancer cases in terms of absolute number of cases among males would range from 1.5 lakhs to 2 lakhs averaging about 1.8 lakhs cases among based on the incidence data of five cancer registries in Urban India which the health planners would need to address for treatment facilities. In females the estimates of absolute number of cases ranged from 1.7 lakhs to 2.2 lakhs based on data from Mumbai and Bengaluru registries respectively. The computations have also been carried for load of tobacco related cancers which this country is likely to face. These computations would help in the program implementation in prioritizing the resources available for health sector.

2. **Applied Statistics 3**

**Paper Name:** An Assessment of Knowledge on Descriptive Statistics among the Postgraduate Students in Sikkim Manipal Institute of Medical Sciences, Sikkim  
**Author:** Rajkumari Sanatombi Devi

**Abstract:**

**INTRODUCTION:**  
Anyone who is involved in medical research should always keep in mind that science is a search for the truth and that, in searching for the truth, there is no room for bias or inaccuracy in statistical analysis and interpretation of data.1 Any errors in statistical analysis will mean that the conclusions of the study may be incorrect.2  
**OBJECTIVE OF THE STUDY**  
The objectives of this study are to assess the knowledge in understanding the
concepts, and identification of appropriate statistical methods among the Post Graduate (PG) students of Sikkim Manipal Institutes of Medical Sciences, Gangtok, Sikkim. The finding of the study will help about the identification of the topics to be developed in their PG curriculum so that they will be able to become an independent and a successful researcher in their profession.

MATERIALS AND METHODS
Study Design: Descriptive study
Study period: July 2013 to July 2015
Study setting: Sikkim Manipal Institute of Medical Sciences, Sikkim
Study population: All the PG (MD/MS/DNB) students enrolled in the First, Second and Third years during the academic session July 2013 to July 2015 are included in the study.

Data collection tool: A questionnaire developed by the researcher is used for data collection. It consists of the common topics which are taught at undergraduate (MBBS) level in medical colleges in India. There are 38 questions of which 18 questions are related from Descriptive Statistics and 20 are based on Inferential Statistics. All the respondents are instructed to choose the correct answers out of 4 options for each question.

Data collection procedure:
After getting the permission from the institute ethical committee, the researcher contacts all the HOD and fix the date and time for the data collection in their departments. The purpose of the study and how to choose the correct answer in the questionnaire are explained by the researcher. Verbal consent is taken from them. Students not willing to involve in the study are excluded from the study. Thirty minutes are given to complete the questionnaire. After that, students are asked to handover the completed questionnaire directly to the researcher.

Sampling technique: Census method will be used for data collection.
Statistical Analysis: Proportion and percentage generated by SPSS 16.0 version for the selected variables will be presented as the results of the study.

Key word: Biostatistics, Knowledge, Questionnaire, Census

**Paper Name:** Systematic Analysis for Community Based Data—A Statistical Modern Methods Approach.

**Author:** Dr. Nandkishore Singh

**Abstract:** In this paper the ideas that Community data to analysis systematically and statistically by surveyor or investigator. Usually we do some mistakes for compile to complete a survey analysis task purposefully. There is one type of sampling unit involved and hierarchical with e.g. communities, households and individuals in Stage of Analysis. To separate out three stages of survey data handling – exploration, analysis and archiving – which help to define expectations
Paper Name: Multi-level Modelling Approach to Analyze RCH Services, Utilization and its Correlates

Author: Tulsi Adhikari

Abstract: The present study aims at emphasizing the role of hierarchical modeling in analyzing the association between the utilization of RCH services covering both maternal and child health related services and its association with the individual level, Community level and District level factors.

The specific objectives of the study are:

To assess the inequity in the access and utilization of different RCH services, viz, maternal (ANC and PNC) and child health care (Immunization and treatment seeking behaviour).

To investigate the degree to which the RCH utilization is influenced by the contexts within which the people live and other explanatory variables and to assess the superiority of Multi-level modeling approach over the Standard Logistic Regression for the current situation.

The DLHS-3 survey data was used to carry out the study. In our study the outcome indicators and the explanatory variables were as under

Outcome indicators (Dependent Variables)

Maternal health

Explanatory Variables (Independent Variable)
District Level Factors

Village Level Factors

Household level factors

Individual (women) Factors

The utilization RCH Indicators under study were found to be associated with the explanatory variables at different level. Further, Comparison of multilevel regression model with the traditional regression mode, shows a strong advocacy for the multilevel modeling approach.

**Paper Name:** Effect of Misallocation in Training Samples on Ability of Discrimination of Multiple Group Discriminant Function (three Group MDF): A Monte Carlo Simulation study.

**Author:** V. S. Yadav

**Abstract:**

Object: In the discriminant analysis the basic assumptions are, (i) Sample come from morally distributed populations (ii) the covariance matrices are same for all populations (iii) the observations in the samples are correctly allocated to their original group. In practice population parameters are unknown. We use their estimates to construct discriminant function. We draw samples and calculate sample mean and covariance. Under above assumptions we perform discriminant analysis using Fisher’s approach. But in practical situations such populations do not exists even though we use robust discriminant function, while estimates are used in construction of linear discriminant function are not robust.

Material & Methods: An attempt is made to evaluate the ability of Fisher’s linear Discriminant Function when training samples are containing misallocation. In the present study following situations are studied (1) - some of the observation in 1st group come from 2nd group and no contamination in second and third group, (2) - some of the observation in 2nd group come from 1st group and no contamination in first and third group, (3) - 1st and 2nd both groups contains misallocated observations from each other and no contamination in third group. Monte Carlo simulation technique is applied in simulation experiments. Two types of sample designs are studied in presence of ten types of contamination in initial samples.

Results: The result shows that linear discriminant function shows robustness up to certain extent of contamination in training sets but higher percentage of misallocation seriously affect the ability of discriminant function. The results of this study will be presented in detail.
**Paper Name:** Crime against women in India: a geospatial analysis

**Author:** Binu V S, Amitha Puranik, Sintomon Mathew, Sebin Thomas

**Abstract:** Globally, women are more vulnerable to various forms of crimes than males. The crimes that are directed specifically towards women are classified as crime against women. Crime against women in India is observed to increase year after year and according to the National Crime Records Bureau (NCRB) report, in 2014 there was an increase of 9.2% cases of crime against women compared to the previous year. The violence in a population depends on socio-demographic factors, unemployment, poverty, number of police officials etc. There are very few studies that explored to identify hot spots of various types of crime against women in India. Hot spots are geographical regions where the number of observed cases is more than the expected number for that region. It is important to identify the hot spots of crime against women in India in order to control and prevent violence against women in that region. The goal of this study is to identify the hotspots of crime against women in India using spatial data analysis techniques. For the present study, we use the district level data of various types of crime against women in India in the year 2011 published by NCRB and the 2011 Census population in each of these districts. The study used Moran’s scatter plot and spatial scan statistic to identify the hotspots using ArcGIS and SaTScan software’s.

**Paper Name:** Factor Analysis and Its Application on Studying the Attachment of Children to Their Parents

**Authors:** K.R Sundaram, Gitanjali M, and Renjumol V.L.

**Abstract:**

Factor analysis is a statistical technique, the aim of which is to simplify a complex data set by representing the set of variables in terms of a smaller number of underlying variables, known as factors or latent variables. The technique is a branch of multivariate analysis in which observed variables are modelled as linear combinations of the factors, plus “error” terms. The method is extensively used in many fields of research including clinical psychology.

In this presentation Factor analysis is applied to the data on attachment of children to their mothers and fathers. The data is taken from a study conducted in Kochi for studying the rejection sensitivity in early adolescence which includes attachment security of children towards mother and father. The questionnaire used for the study
was the Inventory of Parents and Peer Attachment (IPPA). This contained a total of 25 questions with a five point scale. The study sample consisted of 436 school-going children in the age group 12 & 14 years from Ernakulam district. In case of Mothers, five factors were extracted on varimax rotation with a total variation explained as 49.3%. Factor 1 which accounted for 27.1% was Communication (7 items). The other factors were Trust (7.5%-6 items), Alienation (5.5%-7 items), Acceptance of Mother (5.1%-4 items) and Trust & Acceptance (4.1%-1 item). The corresponding variation accounted by these factors in the rotated structure were 11.8%, 10.3%, 10.1%, 8.6% and 8.5%. Results obtained after rotation indicate the relative importance of the factors rather than their independent contribution. In case of Fathers, five factors were extracted with a total variation of 49.6%. Factor 1 which accounted for 28.4% was Communication (9 items). The other factors were Trust (4.7%-5 items), Alienation (2.8%-7 items), Acceptance of Father (1.9%-2 items) and availability (1.3%-2 items). The corresponding variation accounted by these factors in the rotated structure were 15.8%, 12.5%, 9.8%, 6.7% and 4.8%. The results indicate some differences in the nature of attachment of children to mother and father.

3. Health Statistics 4

**Paper Name:** The translation and validation of CAMI (Community Attitude to Mental Illness) questionnaire to Bengali

**Author:** Rivu Basu

**Abstract:**

Introduction: Public stigma against mental health problems is damaging to people with mental illness and is associated with significant societal burden. It is a global phenomenon which is prevalent over time and place. Inspite of growing concern regarding mental health in India, this kind of study to assess the Mental Health Status has been very few in India, more so in West Bengal. To use a standard questionnaire CAMI, it has to be translated first and then validated in Bengali. The study was done to validate this questionnaire.

Methodology: It was an Observational Descriptive study with cross-sectional design done at Amdanga Community Development Block, North 24 Paraganas, West...
Bengal, India. It is the rural field practice area of the Department of Community Medicine, R. G. Kar Medical College and Hospital, Kolkata. Data were collected from 1st December 2015 to 31st January 2016. Adult population (>18 years) of 81 villages of the Amdanga Community Development block were the study population. The schedule that was applied stays valid for adult populations only.

Attitudes to Mental Illness questionnaire was developed by the Department of Health, United Kingdom for this series of surveys, based on previous research in Toronto, Canada and the West Midlands, UK. It included 26 items based on the 40-item Community Attitudes toward the Mentally Ill (CAMI) scale and the Opinions about Mental Illness scale, and an added item on employment-related attitudes. The questions covered a wide range of issues, from attitudes towards people with mental illness, to opinions on services provided for people with mental health problems. The questionnaire was translated into Bengali and Hindi, and again retranslated and checked for consistency. A group of experts of Community Medicine and Psychology Department, Rajabazar Science College tested the judgmental validity. There were 5 interviewers chosen to work in two different teams. They were trained to ensure reliability. A work plan was developed to cover the block in around 2 months. Supervisors from the Department of Community Medicine visited the sites regularly to check progress and assure quality. Data were assembled in Microsoft Excel 2010 software. Other measures of validity were ascertained and factor loading was done to assess the domains. Cronbachs alpha was also calculated to see internal consistency.

Results: The results showed a few questions can be eliminated, and Cronbachs alpha showed good consistency.

Conclusion: The translated version can be safely used to determine Attitude towards mental illness among adult rural population of West Bengal.

**Paper Name:** Timing of menopause: a potential risk for women of reproductive age.

**Author:** AK Mathur, AK Mukherjee, KK Ganguly, RS Sharma

**Abstract:**

Introduction: Menopause is an important transition in women’s reproductive life, as it signals end of fertility. Timings of menopause (natural) is an important determinant of future risk diseases. Studies done in the past have demonstrated that early age of natural menopause (ANM) is associated with increased risk of CVD, osteoporosis where as late menopause increases the risk of breast and endometriosis cancer. Women who have their ovaries removed surgically seems to have less glucose and insulin levels which is suggestive of a link between premature menopause and
diabetes. However past epidemiological studies yielded conflicting results with either inverse or no association with diabetes.

Methods: A cross sectional study adopting multistage random sampling was used in recruiting 3081 women aged above 30 years. Data analysis has been done by stratifying the ANM in three quartiles and exploring its association with factors that may influence the onset of menopause.

Results: overall 1067 (34.6%) women attained natural menopause and the mean and median age at natural menopause (ANM) was found to be 45.7± 5.1 years and 46 years respectively. Data was analyzed using both univariate and multivariate analysis..and indicated that presence of type 2 diabetes was found to be statistically significantly after controlling for education, occupation, intake of tea/coffee, parity and other variables.

Conclusion: Menopausal health demands priority in Indian scenario due to increase in life expectancy and growing population of menopausal women. Concerted efforts are required to educate and make these women aware of menopausal symptoms, reduction of discomfort and enable them to seek appropriate medical care.

Paper Name: Trends in new filarial infections in rural Kanpur – Indications on Mass Drug Administrations

Author: Dr. Anil Kumar, AshisBajpei, T Singh and S Masih

Abstract: Lymphatic filariasis is still a health problem in tropical world including India. Government data (2015) reveals that the disease is still prevalent in 20 states or Union Territory in India and about 1% of the population are microfilaria carriers in India. Globally supported filariasis control programme is functional throughout India for decades and revised national target is set to wipe out the disease by 2020.

One of the important component of the control strategy is to provide annual mass drug administration (MDA) with DEC with/without Albendazol. It was expected that if MDA is repeated for five years, the infection should be functionally controlled or reduce to level below the transmission capabilities. However, studies have reported vary wide variations throughout the globe.

In Uttar Pradesh, several districts are still endemic for disease. In 2004, JALMA adopted around 25 villages in Ghatampur Tehsil of Kanpur District. The initial mf
prevalence was around 12% (8% to 15%) and disease was prevalent in about 6% population of the project area.
In all villages, we undertook mass drug administration using DEC (Diethyl carbamazine citrate 6 mg/kg BW – max 300 mg) alongwith Albendazol -400mg and in some villages provided additional doze 12 days to mf carriers. The strategy was adopted to see how fast sustainable mf clearance can be achieved.
The trend of microfilaria rate, new infection and new disease incidence and also mass drug coverage was monitored. The results reveal that over the years complete drug coverage cannot be achieved and so the clearance of mf rate. This lead to a situation of new cases occurring continuously and therefore suggest that in view of the goal to achieve complete wipe out 5 years strategy could lead to partial success. The details would be presented.

**Paper Name**: Potential of Cognitive Activities in Delaying Dementia among 60 plus Population  

**Abstract**: Dementia is a disease commonly affecting the older population. It leads to decline in mental capabilities, affecting memory and its functioning. Proximal factors also contribute to reduced cognitive performance with aging. Cognitive capacity remains intact with aging, but encoding, storage and retrieval becomes less efficient or are interrupted by reduced attention and working memory capacity. With advancing age, the capacity of the working memory system to hold and manipulate information declines. This age related decline in working memory capacity has implications for higher level aspects of cognition. In various medical studies, it has been observed that no medication has proven to cure dementia. On the contrary these drugs have various adverse effects. This project studies the impact of cognitive activity on working memory. In the present study, workshops of 21 days each were conducted at four old age homes in various parts of Delhi. A total sample population of 96 older adults were part of these workshops. In each workshop MMSE1 and MMSE2 were done at the beginning and at the end of the workshop respectively. Another MMSE3 test was done one month after the completion of the workshop, to access the difference between MMSE2 and 3, when the inmates were not cognitively active. Each session of the workshop consisted of two parts. The first part included entertaining activities which increased the comfort level of the inmates and motivated them to attend the workshop regularly. The second part was a test based on verbal episodic memory or on inductive reasoning. The results found have been encouraging. A significant number of participants showed a remarkable
improvement in cognitive abilities. On comparing MMSE2 and MMSE3 (no cognitive activity) the results showed a drop of 50% in the cognitive abilities.

**Paper Name:** Deriving single pooled summary measure from varied study designs in systematic reviews of Nutritional interventions: A Bayesian approach

**Author:** N. Sreekumaran Nair

**Abstract:**

Randomised Controlled trials are considered to be the gold standard in assessing the effectiveness of clinical interventions. However the same need not get accepted for interventions delivered at population level like nutritional interventions. Studies based on other designs like Non randomised controlled trials, Cohort designs, Quasi randomised trials, programme evaluation etc. been extensively used to evaluate effectiveness of population level intervention programmes. This has been predominantly seen in public health interventions like food fortification using vitamins, minerals and other micronutrient to improve the nutritional status of children, adults or specifically diseased population. Summary evidence from systematic reviews, usually expressed as a single effect measure from many studies, have been frequently used as a powerful information for policy framing of such interventions. However when evidence evaluated by different designs, pooling together these effect measures and expressing as a single effect measure is an issue. Usual practice is to compute summary effect measures based on each design separately and report which defeat the whole idea of meta-analysis. More over this approach again confuse the policy makers who are looking for single effect estimate. In this paper we are proposing a method based on Bayesian approach to get a single summary estimate by incorporating effect measures coming from variety of study designs. The idea is to build a prior using evidence from weaker designs and likelihood from stronger designs. Further use the Bayesian principle to derive the posterior distribution. Ultimately get the pooled effect estimate and its credible interval from the posterior. Authors have applied this method in couple of Nutritional related systematic reviews and obtained promising results. The paper made a comparison of traditional meta-analysis results with the new proposed approach.