Post Graduate Diploma in Biostatistics

Under Regulations- (R-2025)

(w.e.f. 2025-26 admitted batch)

Course Structure and Syllabus

THE APOLLO UNIVERSITY

MURUKAMBATTU - CHITTOOR (Dt) 517127

ANDHRA PRADESH

I. Post Graduate Diploma in Biostatistics

First Semester						
Course Code	**Course Name**	**Periods per week**			**Credits**	**Hours per week**
		L	**T**	**P**		
26BST6501	Fundamentals of Biostatistics	2	2	0	4	4
26BST6502	Statistical Inference	2	2	0	4	4
26BST6503	Design and Analysis of Clinical Trials	2	2	0	4	4
26BST6504 26BST6505	Elective –I (Students will choose one of these Course) i. Data Management and Ethics in Biostatistics Or ii. Health Informatics and Data Science	1	2	0	3	3
Practical/Others						
26BSL6501	Survival Analysis	0	0	2	2	4
26BSL6502	Regression Analysis	0	0	2	2	4
--	Seminar	0	0	0	0	6
--	Library	0	0	0	0	1
--	Extra-curricular activities	0	0	0	0	1
--	Self-Learning/ Out Reach activities	0	0	0	0	5
TOTAL		**7**	**8**	**4**	**19**	**36**

Second Semester						
Course Code	**Course Name**	**Periods per week**			**Credits**	**Hours per week**
		L	**T**	**P**		
26BST6506	Systematic Review	2	2	0	4	4
26BST6507 & 26BST6508	Elective –II (Students will choose one of these Course) i. Epidemiology and Public Health Or ii. Health Economics and Outcomes Research	1	2	0	3	3
Practical/Others						
26BSL6503	Machine Learning for Biostatistics	0	0	2	2	4
26BSI65 01	Dissertation	0	0	12	12	24
TOTAL		**3**	**4**	**14**	**21**	**36**
Grand Total Credits		**10**	**12**	**18**	**40**	

26BST65 01: Fundamentals of Biostatistics

L T P C: 2 2 0 4

Course Description: This course introduces students to the fundamental principles of biostatistics, emphasizing descriptive statistics, probability and different types of data. It equips students with statistical literacy and the ability to apply statistical techniques for data analysis and interpretation.

Course Objectives:

- To introduce students to the basic concepts of biostatistics.
- To equip students with skills to analyse and interpret descriptive statistical measures.
- To understand the fundamentals of probability and its applications in biostatistics.
- To differentiate and categorize different types of data used in statistical analysis.

Unit 1: Introduction to Biostatistics (10 hours)

Definition, scope and importance of biostatistics. Applications in public health and medical research. Types of data including qualitative vs. quantitative, continuous vs. categorical. Data collection techniques and sources, ensuring accuracy and reliability in statistical measurements.

Unit 2: Descriptive Statistics (14 hours)

Measures of central tendency including mean, median and mode. Measures of dispersion such as range, variance and standard deviation. Graphical representation of data using histograms, boxplots and scatter plots. Summarizing data using tables and charts to facilitate interpretation and decision-making.

Unit 3: Probability and Probability Distributions (12 hours)

Basic probability concepts and fundamental rules of probability. Conditional probability and the concept of independence in statistical events. Probability distributions including Binomial, Poisson and Normal distributions. Understanding the Law of Large Numbers and the Central Limit Theorem and their implications in statistical analysis.

Unit 4: Sampling and Sampling Techniques (12 hours)

Concept of sampling and its critical role in statistical analysis. Various types of sampling methods including random, stratified, cluster and systematic sampling. Understanding sampling distribution and standard error. Determination of sample size for accurate statistical inferences and minimizing errors.

Unit 5: Introduction to Statistical Software (12 hours)

Overview of popular statistical software such as SPSS, R and SAS. Data entry and management techniques for large datasets. Basic statistical operations and graphical representation of data using software tools. Interpretation of software-generated outputs for practical applications in biostatistics.

Course Outcomes / Competencies: By the end of this course, students will be able to:

- Apply basic Bio-statistical concepts.
- Perform descriptive statistical analysis and interpret results.
- Apply probability rules in Bio-statistical contexts.
- Utilize statistical software for data analysis.

Textbooks:

- Rosner, B. (2015). Fundamentals of Biostatistics. Cengage Learning.
- Pandey, A. (2019). Biostatistics and Research Methodology. Jaypee Publications.

References:

- Armitage, P., & Berry, G. (2001). Statistical Methods in Medical Research. Wiley.
- Indrayan, A. (2012). Medical Biostatistics. CRC Press.

26BST65 02: Statistical Inference

L T P C: 2 2 0 4

Course Description:

This course focuses on the fundamental concepts and techniques in statistical inference, including hypothesis testing, confidence intervals and error types. Students will gain the ability to make data-driven decisions through rigorous statistical reasoning and develop skills in testing hypotheses and estimating population parameters.

Course Objectives:

- To introduce students to the concepts of hypothesis testing and confidence intervals.
- To develop an understanding of error types in statistical analysis.
- To equip students with the tools necessary for drawing conclusions from sample data and making inferences about populations.

Unit 1: Introduction to Statistical Inference (10 hours)

Definition and importance of statistical inference. Relationship between sample statistics and population parameters. Overview of hypothesis testing and confidence intervals. Basic concepts of estimation and confidence in statistical analysis.

Unit 2: Hypothesis Testing (12 hours)

Null and alternative hypotheses, Type I and Type II errors. Steps in hypothesis testing: formulating hypotheses, selecting significance level, test statistic, p-values and conclusion. One-tailed and two-tailed tests. Common tests: Z-test, t-test, chi-square test and F-test. Interpretation and significance of results in hypothesis testing.

Unit 3: Confidence Intervals (14 hours)

Concept and purpose of a confidence interval. Estimating population parameters and margin of error. Calculation and interpretation of confidence intervals for means, proportions and differences between groups. The relationship between hypothesis testing and confidence intervals. Practical applications in making statistical decisions.

Unit 4: Error Types in Hypothesis Testing (10 hours)

Understanding Type I error (false positive) and Type II error (false negative). The power of a test and factors affecting it. Techniques for minimizing errors in hypothesis testing. Trade-offs between Type I and Type II errors in decision-making.

Unit 5: Practical Applications and Software Tools (14 hours)

Using statistical software (e.g.R, SPSS) for hypothesis testing and confidence interval estimation. Practical examples and case studies in health, business and social sciences. Interpretation of software-generated outputs and error diagnostics in statistical analysis.

Course Outcomes / Competencies for Both Courses:

By the end of the **Statistical Inference** course, students will be able to:

- Conduct hypothesis testing and interpret results.
- Estimate population parameters using confidence intervals.
- Understand and mitigate error types in statistical analysis.
- Apply statistical inference techniques to existing datasets.

Textbooks:

- Howell, D. C. (2012). Statistical Methods for Psychology. Cengage Learning.
- Agresti, A., & Finlay, B. (2009). Statistical Methods for the Social Sciences. Pearson.

References:

- Hogg, R. V., & Tanis, E. A. (2014). Probability and Statistical Inference. Pearson.
- Rice, J. A. (2006). Mathematical Statistics and Data Analysis. Cengage Learning.

26BST65 03: Design and Analysis of Clinical Trials

L T P C: 2 2 0 4

Course Description: This course provides a comprehensive understanding of the design and analysis of clinical trials, covering randomization, blinding, statistical methods and ethical considerations in clinical research.

Course Objectives:

- To introduce the fundamental concepts of clinical trial design.
- To explain the role of randomization and blinding in clinical trials.
- To analyse data from clinical trials using appropriate statistical methods.
- To understand the regulatory and ethical aspects of clinical research.

Unit 1: Introduction to Clinical Trials (10 hours)

Definition and phases of clinical trials. Key components of trial design, including control groups, interventions and endpoints. Ethical considerations in clinical research, including informed consent and patient safety. Overview of regulatory frameworks such as ICH-GCP, FDA, EMA and CDSCO guidelines. Importance of trial registration and data transparency in medical research.

Unit 2: Randomization and Blinding (12 hours)

Importance of randomization in reducing bias. Different randomization methods: simple, stratified, block and adaptive randomization. Advantages and limitations of each method. Blinding techniques: single, double and triple blinding. Impact of blinding on bias reduction and trial validity. Practical implementation of randomization and blinding in clinical trials using statistical software.

Unit 3: Sample Size Determination and Power Analysis (12 hours)

Concepts of statistical power, effect size and significance level in clinical trials. Methods for calculating sample size in different study designs (parallel, crossover, factorial). Power analysis for detecting treatment effects. Interim analysis and stopping rules in clinical trials. Use of statistical software (R, G*Power, SAS) for sample size and power calculations.

Unit 4: Statistical Methods in Clinical Trials (12 hours)

Comparison of Intention-to-Treat (ITT) vs. Per-Protocol (PP) analysis. Statistical methods for analysing categorical (chi-square test, logistic regression) and continuous (t-tests, ANOVA, linear regression) outcomes in clinical trials. Introduction to survival analysis techniques in clinical research. Dealing with missing data: imputation methods and sensitivity analysis. Bayesian vs. frequentist approaches in clinical trial statistics.

Unit 5: Reporting and Interpretation of Clinical Trials (14 hours)

Guidelines for reporting trial results: CONSORT, STROBE, PRISMA. Interpretation of clinical trial outcomes: statistical vs. clinical significance. Conducting and interpreting meta-

analyses and systematic reviews. Use of forest plots, funnel plots and sensitivity analyses in evidence synthesis. Practical applications using statistical software for summarizing and visualizing trial results.

Course Outcomes / Competencies: By the end of this course, students will be able to:

- Design and implement clinical trials following best practices.
- Apply appropriate randomization and blinding methods.
- Conduct statistical analyses for clinical trial data.
- Interpret and communicate clinical trial findings.

Textbooks:

- Friedman, L.M., Furberg, C.D., & DeMets, D.L. (2010). Fundamentals of Clinical Trials. Springer.
- Pandya, N. (2020). Essentials of Clinical Research. Jaypee Publications.

References:

- Gøtzsche, P.C. (2011). Rethinking Clinical Trials: A Design and Analysis Perspective. Wiley.
- Machin, D., Day, S., & Green, S. (2006). Textbook of Clinical Trials. Wiley.

Elective Courses: I

26BST65 04: Data Management and Ethics in Biostatistics

L T P C: 1 2 0 3

Course Description:

This course introduces students to the principles and practices of data management and ethics in biostatistics. It covers the importance of data quality, data cleaning and the ethical considerations involved in collecting, storing and analysing data. Students will also learn about maintaining data integrity, handling missing data and ensuring compliance with ethical standards in Bio-statistical research.

Course Objectives:

- To familiarize students with data management processes in biostatistics, including data collection, cleaning and validation.
- To introduce ethical issues related to Bio-statistical research and data handling.
- To enable students to apply ethical guidelines and best practices in managing and analysing data.

Unit 1: Introduction to Data Management (8 hours)

Overview of data management in biostatistics. Data types, sources and collection methods. Data storage and organization: databases, spreadsheets and data management systems. Principles of data quality: accuracy, completeness, consistency and timeliness. Understanding the data life cycle from collection to analysis.

Unit 2: Data Cleaning and Validation (9 hours)

Techniques for cleaning data: identifying and correcting errors, handling outliers and addressing inconsistencies. Validation of data: checking for missing values, duplicates and logical errors. Data transformation and normalization. Use of software tools (e.g.R, SAS) for data cleaning and validation. Ensuring reproducibility and reliability of data.

Unit 3: Ethical Considerations in Biostatistics (10 hours)

Ethical principles in biostatistics: respect for persons, beneficence and justice. The role of ethics in data collection, storage and analysis. Informed consent in data collection: privacy, confidentiality and data protection. Ethical issues in the use of secondary data and sharing of research findings. Compliance with regulatory standards (e.g.IRB approval, HIPAA, GDPR).

Unit 4: Handling Missing Data (10 hours)

Types of missing data: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). Techniques for handling missing data: imputation methods (mean imputation, regression imputation, multiple imputation) and sensitivity analysis. The

impact of missing data on statistical analysis and research findings. Best practices for reporting and handling missing data in Bio-statistical studies.

Unit 5: Data Integrity and Ethical Decision-Making (8 hours)

Ensuring data integrity: preventing data manipulation and fraud. The role of transparency and accountability in Bio-statistical research. Ethical decision-making in data analysis: balancing scientific integrity with research goals. Case studies of ethical dilemmas in biostatistics. Professional and legal responsibilities of biostatisticians.

Course Outcomes / Competencies:

By the end of the **Data Management and Ethics in Biostatistics** course, students will be able to:

- Understand the principles and best practices of data management in biostatistics.
- Apply data cleaning and validation techniques to ensure high-quality datasets.
- Navigate ethical issues in data collection, analysis and reporting.
- Handle missing data effectively and apply ethical decision-making in Bio-statistical research.

Textbooks:

- Khatri, C., & Mahajan, K. (2013). Data Management in Biostatistics: Concepts and Practices. Springer.
- Vickers, A. J., & Altman, D. G. (2001). Statistics in Medicine. Wiley.

References:

- Horton, N. J., & Kleinman, K. P. (2015). Data Management and Ethics in Biostatistics: Principles and Techniques. Springer.
- Altman, D. G., & Bland, J. M. (2009). Missing Data. BMJ.

26BST65 05: Health Informatics and Data Science

L T P C: 1 2 0 3

Course Description: This course covers the role of data science in healthcare, including data collection, analysis, visualization and the use of informatics in decision-making.

Course Objectives:

- To introduce students to the principles of health informatics and data science.
- To explore various methods of health data analysis and visualization.
- To understand electronic health records and decision support systems.

Unit 1: Introduction to Health Informatics (8 hours)

Definition and scope of health informatics. Evolution and applications of health informatics in public health and clinical settings. Key components of health information systems. Health data standards and interoperability (HL7, FHIR, DICOM). Electronic Health Records (EHR): structure, functionalities and challenges in implementation. Decision support systems and their role in healthcare management.

Unit 2: Data Science in Healthcare (8 hours)

Types and sources of healthcare data: administrative data, clinical data, biomedical data and wearable technology data. Structured vs. unstructured data in healthcare. Data pre-processing and data cleaning techniques for health informatics applications. Introduction to predictive analytics and its role in clinical decision-making. Case studies on predictive modelling in healthcare.

Unit 3: Visualization and Interpretation of Health Data (12 hours)

Fundamentals of data visualization in health informatics. Common visualization tools and techniques: dashboards, heatmaps and geographic information systems (GIS). Role of visual analytics in healthcare decision-making. Case studies on the application of visualization in hospital management, disease surveillance and public health monitoring. Implementation of dashboards for real-time monitoring of patient data.

Unit 4: Artificial Intelligence and Machine Learning in Healthcare (9 hours)

Introduction to artificial intelligence and its applications in healthcare. Overview of machine learning algorithms used in healthcare analytics: decision trees, support vector machines, neural networks and deep learning. AI-driven medical diagnosis, disease prediction and personalized medicine. Challenges and limitations of AI in healthcare. Ethical considerations in AI-driven decision-making and patient care.

Unit 5: Ethical and Legal Issues in Health Informatics (8 hours)

Principles of data privacy and security in health informatics. Legal frameworks and regulations: HIPAA (Health Insurance Portability and Accountability Act), GDPR (General Data Protection Regulation) and other international standards. Ethical concerns in the use of big data

and AI in healthcare. Challenges in maintaining patient confidentiality and data security. Case studies on data breaches and ethical dilemmas in health informatics.

Course Outcomes / Competencies: By the end of this course, students will be able to:

- Understand the role of data science in healthcare decision-making.
- Analyze and visualize health data effectively.
- Apply machine learning techniques in healthcare scenarios.
- Navigate ethical and legal considerations in health informatics.

Textbooks:

- Shortliffe, E.H., & Cimino, J.J. (2014). Biomedical Informatics: Computer Applications in Health Care. Springer.
- Bansal, M. (2019). Health Informatics: An Overview. Jaypee Publications.

Reference Books

- Hoyt, R.E., & Yoshihashi, A. (2020). Health Informatics: Practical Guide. Lulu.com.
- Ristevski, B., & Chen, M. (2018). Big Data Analytics in Healthcare. Springer.
- Dey, N., & Ashour, A.S. (2019). Machine Learning in Healthcare Informatics. Springer.

26BST65 01: Practical Survival Analysis

L T P C: 0 0 2 4

Course Description:

This course provides hands-on training in survival analysis techniques with a focus on real-world applications in public health and clinical research. Students will analyze time-to-event data using statistical software (R, SPSS, jamovi, SAS), construct Kaplan-Meier survival curves, apply Cox proportional hazards models and perform model diagnostics. Emphasis is placed on interpreting results, assessing model assumptions and applying survival analysis methods to clinical and epidemiological datasets.

Course Objectives:

- To develop proficiency in handling survival data using statistical software (R, SPSS, jamovi, SAS).
- To apply Kaplan-Meier and Cox proportional hazards models to real-world datasets.
- To perform model diagnostics and extensions for better interpretation of survival data.
- To analyze and interpret survival data from clinical trials and epidemiological studies.

Unit 1: Introduction to Survival Analysis (8 hours)

Practical Focus:

- Overview of survival data: Types of censoring (right-censored, left-censored, interval-censored).
- Hands-on session: Importing and exploring survival datasets in R/SPSS/jamovi/SAS.
- Constructing survival and hazard functions using real-world datasets.

Activities:

- Explore publicly available survival datasets (e.g.SEER cancer data, cardiovascular studies).
- Visualize survival and hazard functions using software tools.

Unit 2: Kaplan-Meier Estimator (12 hours)

Practical Focus:

- Step-by-step construction of Kaplan-Meier survival curves.
- Performing and interpreting log-rank tests for comparing survival curves.
- Practical exercises using clinical datasets (e.g.oncology, infectious diseases).

Activities:

- Analyze survival data from clinical trials to estimate survival probabilities.

- Compare survival curves between groups using log-rank tests.
- Interpret results and present findings in graphical formats.

Unit 3: Cox Proportional Hazards Model **(12 hours)**

Practical Focus:

- Fitting a Cox proportional hazards model using R/SPSS/jamovi/SAS.
- Understanding hazard ratios and their interpretations.
- Model evaluation and significance testing.

Activities:

- Fit Cox models to real-world datasets (e.g.cardiovascular risk factors, cancer recurrence).
- Interpret hazard ratios and assess the impact of covariates on survival outcomes.
- Evaluate model fit and test for significance of predictors.

Unit 4: Model Diagnostics and Extensions **(14 hours)**

Practical Focus:

- Checking the proportional hazards assumption using graphical and statistical methods.
- Residual analysis and interpretation of diagnostic plots.
- Handling time-dependent covariates, competing risks and multi-state models.

Activities:

- Test the proportional hazards assumption using Schoenfeld residuals.
- Perform residual analysis to identify influential points or outliers.
- Extend models to include time-dependent covariates and competing risks.

Unit 5: Clinical Applications of Survival Analysis **(14 hours)**

Practical Focus:

- Application of survival analysis methods in real-life clinical trials and epidemiological research.
- Case studies in oncology, cardiovascular diseases and infectious diseases.
- Project-based learning: Students analyze a survival dataset and present findings.

Activities:

- Work on case studies involving real-world datasets (e.g.cancer survival, HIV progression).
- Conduct a complete survival analysis project, including data cleaning, model fitting, diagnostics and interpretation.
- Present findings in a structured report format, including tables, graphs and conclusions.

Course Outcomes / Competencies:

By the end of the course, students will be able to:

- Conduct survival analysis using statistical software (R, SPSS, jamovi, SAS).
- Apply Kaplan-Meier and Cox proportional hazards models to real-world datasets.
- Assess model assumptions and interpret diagnostic results effectively.
- Communicate survival analysis findings clearly in a public health or clinical context.

Textbooks:

- Kleinbaum, D. G., & Klein, M. (2012). Survival Analysis: A Self-Learning Text . Springer.
- Hosmer, D. W., & Lemeshow, S. (1999). Applied Survival Analysis: Regression Modeling of Time to Event Data . Wiley.

References:

- Therneau, T. M., & Grambsch, P. M. (2000). Modeling Survival Data: Extending the Cox Model . Springer.
- Cox, D. R. (1972). Regression Models and Life-Tables . Journal of the Royal Statistical Society, Series B.
- Harrell, F. E. (2015). Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression and Survival Analysis . Springer.

26BST65 02: Practical - Regression Analysis

L T P C: 0 0 2 4

Course Description:

This course focuses on the practical application of regression analysis techniques in real-world datasets. Students will develop skills in building, diagnosing and interpreting regression models using statistical software tools. Emphasis will be placed on model validation, prediction and interpretation of results in various fields such as health, economics and social sciences.

Course Objectives:

- To equip students with hands-on skills for developing and interpreting regression models.
- To apply diagnostic techniques to validate regression models and improve their performance.
- To analyze real-world datasets and interpret results for decision-making.
- To use statistical software tools effectively for regression analysis.

Unit 1: Introduction to Regression Analysis in Practice (8 hours)

Practical Focus:

- Overview of regression analysis and its applications.
- Types of regression models: Simple linear regression and multiple linear regression.
- Hands-on introduction to statistical software (e.g.R, Python, SPSS).
- Data preparation: Importing datasets, cleaning data and handling missing values.
- Identifying dependent and independent variables in real-world datasets.

Activities:

- Import and explore datasets (e.g.healthcare, economic or social science data).
- Perform exploratory data analysis (EDA) to understand relationships between variables.

Unit 2: Simple Linear Regression: Hands-On Exercises (14 hours)

Practical Focus:

- Fitting simple linear regression models using software tools.
- Estimation of regression coefficients and their interpretation.
- Hypothesis testing and goodness-of-fit measures (R-squared, adjusted R-squared).
- Residual analysis to check model assumptions.

Activities:

- Fit simple linear regression models to real-world datasets.
- Interpret regression coefficients and assess model fit.

- Conduct hypothesis tests for regression coefficients.
- Visualize residuals and identify potential issues (e.g.non-linearity).

Unit 3: Multiple Linear Regression: Practical Modelling (12 hours)

Practical Focus:

- Extending simple linear regression to multiple predictors.
- Handling multi-collinearity: Detection (VIF) and solutions.
- Estimation of coefficients and significance testing.
- Evaluating model performance using metrics like adjusted R-squared, AIC and BIC.

Activities:

- Build multiple linear regression models for complex datasets.
- Detect and address multi-collinearity using variance inflation factor (VIF).
- Compare models using AIC and BIC.
- Interpret results and draw actionable insights.

Unit 4: Model Diagnostics and Validation (14 hours)

Practical Focus:

- Checking assumptions of linear regression: Linearity, normality and homoscedasticity.
- Residual analysis: Identifying influential points and outliers.
- Transformations to improve model fit (e.g.log transformations).
- Cross-validation techniques for model validation.

Activities:

- Perform residual analysis and diagnose model issues.
- Apply transformations to improve model assumptions.
- Use cross-validation to evaluate model performance on unseen data.
- Validate models using real-world test datasets.

Unit 5: Applications of Regression Analysis (12 hours)

Practical Focus:

- Predictive modeling and forecasting using regression.
- Case studies in health, economics and social sciences.
- Interpreting model outputs and communicating results.
- Ethical considerations in regression analysis (e.g.bias, fairness).

Activities:

- Work on case studies involving real-world datasets (e.g.predicting disease outcomes, economic trends).
- Develop predictive models and evaluate their accuracy.
- Present findings in a structured report format.
- Discuss ethical challenges in regression analysis and propose solutions.

Course Outcomes / Competencies:

By the end of this course, students will be able to:

- Develop and interpret regression models using statistical software.
- Perform diagnostics to ensure model assumptions are met.
- Apply regression analysis to real-world datasets for prediction and inference.
- Communicate results effectively through reports and presentations.

Textbooks:

- Montgomery, D. C., & Peck, E. A. (2001). Introduction to Linear Regression Analysis. Wiley.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). Applied Linear Statistical Models . McGraw-Hill.

References:

- Draper, N. R., & Smith, H. (2014). Applied Regression Analysis . Wiley.
- Weisberg, S. (2005). Applied Linear Regression . Wiley.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning . Springer.

Second Semester

26BST65 06: Systematic Reviews

L T P C: 2 2 0 4

Course Description:

This course focuses exclusively on systematic reviews as a rigorous method for synthesizing evidence in health research. It emphasizes the principles, processes and tools required to conduct high-quality systematic reviews. Students will learn how to formulate research questions, develop protocols, search for and appraise evidence and synthesize findings using standardized methods. The course also introduces open-source software for managing systematic reviews and ensures students gain practical skills in critical appraisal.

Course Objectives:

- To understand the principles and rationale of systematic reviews in evidence-based healthcare.
- To develop skills in planning, conducting and reporting systematic reviews.
- To critically appraise the quality of systematic reviews and their role in decision-making.
- To use open-source software for managing and synthesizing evidence in systematic reviews.

Unit 1: Introduction to Systematic Reviews (10 hours)

Definition, scope and importance of systematic reviews in evidence-based healthcare. Rationale and potential of systematic reviews in synthesizing evidence. Principles and procedures of systematic reviews: transparency, reproducibility and rigor. Types of systematic reviews: Reviews of observational studies, clinical trials, diagnostic studies and qualitative evidence. Differences between narrative reviews and systematic reviews. Advantages, challenges and limitations of conducting systematic reviews.

Unit 2: Planning and Protocol Development (12 hours)

Steps in planning a systematic review: Formulating a clear research question using frameworks like PICO (Population, Intervention, Comparison, Outcome). Writing a protocol: objectives, inclusion/exclusion criteria and methodology. Developing a comprehensive search strategy: Identifying relevant databases (e.g.PubMed, Cochrane Library, Embase). Using Boolean operators, MeSH terms and filters. Tools for managing references: EndNote, Zotero or Mendeley.

Unit 3: Study Selection and Quality Appraisal (14 hours)

Locating and selecting studies based on predefined eligibility criteria. Screening titles, abstracts and full texts using standardized forms. Assessing the quality of included studies:

Tools for risk-of-bias assessment (e.g.Cochrane Risk of Bias Tool, ROBINS-I). Evaluating study validity and reliability. Ensuring transparency and reproducibility in the selection process.

Unit 4: Data Extraction and Synthesis (10 hours)

Extracting data from included studies using standardized templates. Organizing and summarizing evidence: Narrative synthesis of findings. Tabular and graphical presentation of results. Introduction to Review Manager (RevMan) software for systematic reviews. Practical exercises: Using RevMan for data management and visualization.

Unit 5: Critical Appraisal and Reporting (14 hours)

Critical appraisal of systematic reviews: Assessing the quality of existing reviews using tools like AMSTAR-2. Identifying strengths, weaknesses and biases in published reviews. Reporting guidelines for systematic reviews: PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Writing and presenting systematic reviews: Structuring the review (Introduction, Methods, Results, Discussion). Communicating findings to diverse audiences (clinicians, policymakers, researchers). Ethical considerations in systematic reviews: avoiding bias, ensuring transparency and addressing conflicts of interest.

Course Outcomes / Competencies:

By the end of this course, students will be able to:

- Understand the principles, rationale and steps involved in conducting systematic reviews.
- Develop a protocol and execute a systematic review using standardized methods.
- Critically appraise the quality of systematic reviews and interpret their findings.
- Use open-source software (e.g.RevMan) for managing and synthesizing evidence.
- Write and report systematic reviews following PRISMA guidelines.

Textbooks:

- Cochrane Handbook for Systematic Reviews of Interventions by Julian PT Higgins and Sally Green, Oxford: Cochrane Collaboration, 2011.
- Systematic Reviews and Meta-Analysis: A Practical Guide by Julia H. Littell, Jacqueline Corcoran and Vijayan Pillai (Sage Publications, 2008).

Reference Books:

- The Handbook of Research Synthesis by Harris Cooper and Larry V. Hedges, Russell Sage Foundation, 1994.
- Systematic Reviews and Meta-Analysis: Pocket Guide to Social Work Research Methods by Julia H. Littell (Oxford University Press, 2008).
- PRISMA Guidelines for Systematic Reviews and Meta-Analyses (available online).

Elective –II

26BST65 07: Epidemiology and Public Health

L T P C: 1 2 0 3

Course Description:

This course introduces students to the principles of epidemiology and public health, focusing on disease prevention, outbreak investigation and the application of statistical methods in health research. It emphasizes hands-on learning through real-world datasets, case studies and the use of statistical software (e.g.R, Jamovi, SPSS) for epidemiological analysis. Students will develop skills in outbreak investigation, study design and evaluating public health policies.

Course Objectives:

- To understand basic epidemiological concepts and measures of disease frequency.
- To develop skills in outbreak investigation and epidemiological study designs.
- To analyze disease surveillance methods and public health policies using statistical tools.
- To apply ethical considerations in epidemiological research and public health interventions.

Unit 1: Introduction to Epidemiology (8 hours)

Definition, history and scope of epidemiology. Key measures: Incidence, prevalence, mortality rates and case fatality rates. Epidemiological transition and global burden of diseases. Applications of epidemiology in public health and healthcare systems. Practical Focus: Hands-on exercises on calculating incidence and prevalence using real-world datasets. Use of statistical software (R or Jamovi) for basic descriptive epidemiological analysis.

Outcome: Students will be able to define and measure disease occurrence in populations and interpret epidemiological data.

Unit 2: Study Designs in Epidemiology (10 hours)

Observational studies: Cohort, case-control, cross-sectional. Experimental studies: Randomized controlled trials (RCTs). Bias, confounding and strategies to mitigate them. Strengths and limitations of different study designs. Practical Focus: Designing a hypothetical cohort study or case-control study. Identifying bias and confounding in published studies. Hands-on exercise: Analyzing a dataset from an observational study using statistical software.

Outcome: Students will understand different epidemiological study designs and their applications in research.

Unit 3: Disease Surveillance and Outbreak Investigation (12 hours)

Principles and methods of disease surveillance (active vs. passive surveillance). Steps in investigating an outbreak: Case definition, line listing, hypothesis generation and testing. Use of geographic information systems (GIS) in disease mapping. Practical Focus: Case studies on real-world outbreaks (e.g.COVID-19, Ebola, cholera). Hands-on outbreak investigation using simulated datasets. GIS-based visualization of disease spread and hotspots.

Outcome: Students will gain practical skills in disease surveillance and outbreak response.

Unit 4: Epidemiology of Communicable and Non-Communicable Diseases (10 hours)

Epidemiology of infectious diseases (TB, HIV/AIDS, COVID-19). Chronic disease epidemiology (Diabetes, Cardiovascular diseases, Cancer). Prevention and control strategies: Vaccination, screening, lifestyle interventions. Practical Focus: Analyzing trends in communicable and non-communicable diseases using global databases (e.g.WHO Global Health Observatory, CDC).Hands-on exercise: Developing a disease prevention strategy for a specific population.

Outcome: Students will be able to assess and evaluate disease burden and prevention strategies.

Unit 5: Public Health Policies and Interventions (5 hours)

Role of WHO, CDC, ICMR and other agencies in shaping public health policies. Health programs in India (e.g.National Health Mission, Ayushman Bharat) and globally. Ethical considerations in epidemiological research and public health interventions. Practical Focus: Group discussion on the effectiveness of national and international health programs. Case study: Evaluating the impact of a public health policy using statistical tools.

Outcome: Students will understand the role of policies in shaping public health outcomes and ethical challenges in research.

Course Outcomes / Competencies:

By the end of this course, students will be able to:

- Define and measure disease occurrence in populations using epidemiological metrics.
- Design and critically evaluate epidemiological studies, including identifying bias and confounding.
- Conduct disease surveillance and outbreak investigations using real-world datasets and tools like GIS.
- Assess the burden of communicable and non-communicable diseases and propose evidence-based prevention strategies.
- Evaluate public health policies and interventions while considering ethical implications.

Textbooks:

- Gordis, L. (2019). Epidemiology . Elsevier.
- Friis, R. H., & Sellers, T. A. (2021). Epidemiology for Public Health Practice . Jones & Bartlett Learning.

Reference Books:

- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). Modern Epidemiology . Lippincott Williams & Wilkins.
- Beaglehole, R., Bonita, R., & Kjellström, T. (2006). Basic Epidemiology . WHO.
- Detels, R., Gulliford, M., Karim, Q. A., & Tan, C. C. (2015). Oxford Textbook of Global Public Health . Oxford University Press.

26BST65 08: Health Economics and Outcomes Research

L T P C: 1 2 0 3

Course Description:

This course introduces students to the principles and methods of health economics and outcomes research (HEOR). It covers economic evaluation techniques, including cost-effectiveness and cost-utility analysis and the assessment of healthcare outcomes. Students will learn to apply these methods in healthcare decision-making, particularly in the context of policy and clinical practice.

Course Objectives:

- To introduce students to the fundamentals of health economics and its role in healthcare decision-making.
- To equip students with methods for evaluating healthcare interventions using economic and outcomes-based measures.
- To develop the ability to analyse and interpret economic evaluations and outcomes data for healthcare policy and clinical applications.

Unit 1: Introduction to Health Economics (8 hours)

Definition and scope of health economics. The role of economics in healthcare policy and practice. Basic economic concepts: demand, supply and market equilibrium in healthcare. The concept of opportunity cost and efficiency in healthcare. Introduction to the economic evaluation framework in healthcare.

Unit 2: Economic Evaluation Techniques (10 hours)

Types of economic evaluations: cost-effectiveness analysis (CEA), cost-utility analysis (CUA), cost-benefit analysis (CBA) and cost-minimization analysis (CMA). The concept of Quality-Adjusted Life Years (QALYs) and Disability-Adjusted Life Years (DALYs). Methods for calculating and comparing costs and outcomes in healthcare interventions. Practical applications of economic evaluation in healthcare.

Unit 3: Measuring Healthcare Outcomes (10 hours)

Defining and measuring healthcare outcomes: clinical outcomes, quality of life and patient-reported outcomes (PROs). Health status measures: generic vs. disease-specific instruments. Health-related quality of life (HRQoL) assessment tools. The importance of outcome measures in economic evaluations and policy decisions. Methods for collecting and interpreting outcomes data.

Unit 4: Health Technology Assessment (HTA) and Policy (8 hours)

Health Technology Assessment (HTA): definition, methods and role in decision-making. The role of HTA in evaluating new medical technologies, pharmaceuticals and healthcare interventions. Regulatory frameworks and agencies involved in HTA (e.g.NICE, ICER). The use of HTA evidence in healthcare policy and reimbursement decisions.

Unit 5: Applications of HEOR in Healthcare Decision-Making (8 hours)

Applying health economics and outcomes research in clinical settings. Cost-effectiveness in healthcare resource allocation. Evaluating public health interventions, pharmaceuticals and medical devices. Decision modelling in healthcare economics. Case studies in pharmaceutical pricing, healthcare policy development and insurance systems.

Course Outcomes / Competencies:

By the end of the **Health Economics and Outcomes Research** course, students will be able to:

- Understand the key principles and concepts in health economics and outcomes research.
- Apply economic evaluation techniques (e.g.cost-effectiveness, cost-utility) to assess healthcare interventions.
- Measure and interpret healthcare outcomes, including quality of life and patient-reported outcomes.
- Conduct health technology assessments and apply HEOR in policy, clinical and reimbursement decision-making.

Textbooks:

- Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press.
- Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (2011). Cost-Effectiveness in Health and Medicine. Oxford University Press.

References:

- Brouwer, W. B., & Koopmanschap, M. A. (2007). Economic Evaluation in Health Care: A Guide to the Methods and Applications of Economic Evaluation of Health Care Treatments. Oxford University Press.
- Revill, P., & Morton, A. (2019). Health Economics: An Introduction for Health Professionals. Wiley.

26BS6503: Practical - Machine Learning for Biostatistics

L T P C: 0 0 2 4

Course Description:

This course focuses on the practical application of machine learning techniques in biostatistics and healthcare analytics. Students will develop skills in implementing supervised and unsupervised learning algorithms, feature selection, deep learning models and model evaluation using statistical software tools (e.g.R, Python, SPSS, SAS). Emphasis is placed on analyzing real-world biomedical and epidemiological datasets, addressing ethical challenges and deploying AI-driven solutions in healthcare settings.

Course Objectives:

- To equip students with hands-on skills for applying machine learning techniques to biomedical and clinical datasets.
- To implement supervised and unsupervised learning algorithms for disease prediction, patient segmentation and healthcare decision-making.
- To optimize deep learning models for medical imaging and time-series healthcare data.
- To address ethical concerns and regulatory requirements in AI-driven healthcare applications.

Unit 1: Introduction to Machine Learning Tools (8 hours)

Practical Focus:

- Overview of machine learning tools: R, Python (Scikit-Learn, TensorFlow, Keras), SPSS and SAS.
- Data pre-processing techniques: Handling missing data, normalization and feature scaling.
- Importing and exploring biomedical datasets (e.g.clinical trials, electronic health records, genomics).

Activities:

- Install and configure machine learning libraries (e.g.Scikit-Learn, TensorFlow).
- Perform data cleaning and pre-processing on real-world datasets.
- Visualize datasets using libraries like Matplotlib and Seaborn.

Unit 2: Supervised Learning: Hands-On Exercises (12 hours)

Practical Focus:

- Implementing regression models: Linear regression, logistic regression, ridge and lasso regression.
- Classification techniques: Decision trees, random forests, gradient boosting (XGBoost, LightGBM) and Support Vector Machines (SVM).
- Model evaluation metrics: Accuracy, precision, recall, F1-score and ROC-AUC.

Activities:

- Build predictive models for disease risk prediction using supervised learning algorithms.
- Evaluate model performance using cross-validation and confusion matrices.
- Case study: Predicting patient outcomes using clinical trial data.

Unit 3: Unsupervised Learning: Clustering and Dimensionality Reduction (10 hours)

Practical Focus:

- Clustering techniques: K-Means, hierarchical clustering and DBSCAN.
- Dimensionality reduction methods: Principal Component Analysis (PCA) and t-SNE.
- Feature selection and extraction for biomedical data analysis.

Activities:

- Segment patients based on clinical features using clustering algorithms.
- Reduce dimensionality of high-dimensional genomic datasets using PCA and t-SNE.
- Case study: Patient segmentation in epidemiological research.

Unit 4: Deep Learning Models: Medical Imaging and Time-Series Data (10 hours)

Practical Focus:

- Building artificial neural networks (ANNs) and deep learning models.
- Convolutional Neural Networks (CNNs) for medical image analysis (e.g.X-rays, MRI scans).
- Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks for time-series healthcare data.
- Model optimization and hyperparameter tuning.

Activities:

- Train CNNs for disease diagnosis using medical imaging datasets.
- Analyze time-series data (e.g.heart rate monitoring) using RNNs and LSTMs.
- Optimize deep learning models to prevent overfitting.

Unit 5: Ethical Considerations and Model Deployment (5 hours)

Practical Focus:

- Addressing bias and fairness in machine learning models for healthcare.
- Ensuring interpretability and explainability of models in clinical decision-making.
- Deploying machine learning models in healthcare applications.

Activities:

- Evaluate ethical challenges in AI-driven healthcare (e.g.HIPAA, GDPR compliance).
- Deploy a trained model using cloud platforms (e.g.AWS, Azure) or web frameworks (e.g.Flask, Django).
- Group discussion: Future trends in AI-driven drug discovery and personalized medicine.

Course Outcomes / Competencies:

By the end of this course, students will be able to:

- Apply machine learning tools (R, Python, SPSS, SAS) to analyze biomedical and clinical datasets.
- Implement supervised and unsupervised learning algorithms for disease prediction, patient segmentation and healthcare decision-making.
- Develop and optimize deep learning models for medical imaging and time-series healthcare data.
- Evaluate machine learning models using performance metrics and validation techniques.
- Address ethical concerns and deploy AI-driven solutions in healthcare settings.

Textbooks:

- Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow . O'Reilly Media.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning . Springer.

References:

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning . Springer.
- Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective . MIT Press.
- Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine . NEJM AI Journal.

26BSI6501: **Dissertation – Comprehensive Viva-Voce**

L T P C: 0 0 8 16

Course Description:

Biostatistics plays a crucial role in public health research by providing the tools to analyse and interpret data effectively. In the PG Diploma in Biostatistics program, the dissertation is an essential component, providing students the opportunity to apply Bio-statistical methods to existing public health issues. The research undertaken should be of high academic rigor and demonstrate the student's ability to independently apply Bio-statistical techniques in a public health context.

The dissertation process will involve the formulation of a research question, data collection, analysis using Bio-statistical techniques and interpretation of results. Students will defend their dissertation through a viva voce, demonstrating a comprehensive understanding of their research process, results and public health implications.

Course Objectives:

This course aims to equip PG Diploma in Biostatistics students with the necessary research skills to carry out a dissertation that applies statistical methods to public health issues. Students are expected to:

1. **Conduct independent research:** Select a public health topic and apply advanced Biostatistical methods for analysis.
2. **Enhance critical thinking:** Evaluate public health data, choose appropriate methodologies and interpret statistical results.
3. **Apply Biostatistical tools:** Utilize regression models, survival analysis and hypothesis testing in research.
4. **Develop reporting skills:** Communicate statistical findings effectively in written and oral formats.
5. **Prepare for viva voce:** Defend methodology, results and conclusions, demonstrating mastery of biostatistics**.**

Project Process:

1. **Project Allocation and Supervision**:
    - A project supervisor will be assigned to each student. The student may also have 1-2 co-supervisors, preferably with expertise in biostatistics or public health.

2. **Project Topic Selection**:
    - The student, in consultation with the supervisor(s), will select a research topic relevant to biostatistics and public health. Topics may include, but are not limited to:

- Statistical modelling of public health data.
- Epidemiological studies involving Bio-statistical techniques.
- Analysis of health outcomes or risk factors.
- Statistical review of public health policies.
- Systematic reviews with a Bio-statistical perspective.

3. **Project Proposal**:
   - A brief synopsis of the proposed project (including title, background, research objectives, methods and analysis plan) must be submitted to the supervisor.
   - If the project involves human participants, ethical approval and consent documents are required for the study.

4. **Data Collection and Analysis**:
   - Following ethics approval, students will collect data (if applicable) and perform statistical analysis using Bio-statistical software (e.g.R, SPSS, SAS).
   - The research design should be robust and data analysis should be thorough, employing appropriate statistical techniques to answer the research question.

5. **Final Dissertation Report**:
   - The report should follow a formal structure, including:
     - Title page, certificate of fulfilment, table of contents and acknowledgments.
     - Introduction, literature review, research objectives, methodology (with statistical methods described in detail), results (including tables, figures and statistical analysis), discussion, conclusion, limitations and references.
     - The final report should not exceed 60 pages (including appendices and references) and should be written in clear, concise English using 12-point Times New Roman font, 1.5 line spacing.
   - One hard copy and one soft copy of the dissertation must be submitted.

6. **Viva Voce Examination**:
   - The student must defend their dissertation in a viva voce exam, where they will be questioned by a panel of experts in biostatistics and public health. The viva voce will assess:
     - The student's understanding of Bio-statistical methods used.
     - The ability to interpret and explain research results.
     - The application of biostatistics to solve public health issues.

7. **Dissertation Defence**:
   - The viva voce will carry 100% of the assessment weight, with 75% of the marks awarded by a panel of departmental experts and 25% by the supervisor(s).

Course Outcomes / Competencies:

Upon successful completion of the dissertation, students will:

1. Demonstrate proficiency in applying advanced Bio-statistical methods to existing public health problems.
2. Conduct rigorous data analysis and draw valid conclusions from public health data.

3. Develop and present a coherent research proposal, including statistical methods and analysis plans.
4. Synthesize and critically analyse existing literature related to biostatistics and public health.
5. Communicate research findings effectively through both written reports and oral presentations, defending their work in a professional setting.

Textbooks:

- Fundamentals of Biostatistics by Bernard Rosner, Cengage Learning, 2015.
- Statistics for Epidemiology by Nicholas P. Jewell, CRC Press, 2009.

Reference Books:

- Systematic Reviews in Health Care: A Practical Guide by Paul Glasziou, Les Irwig, Chris Bain and Graham Colditz, Cambridge University Press, 2001.
- Applied Biostatistics by Richard F. Stokes, Wiley, 2008.
- The Public Health Researcher's Handbook by Ron Iphofen, Policy Press, 2017.