

**WORKSHOP ON EPIDEMIOLOGICAL
AND BIO-STATISTICAL METHODS
USING 'R' SOFTWARE**

(24th – 29th May, 2010)

COURSE REPORT



आरोग्यम् सुखसम्पदा
NIHFV

**National Institute of Health & Family Welfare
New Delhi**

I. INTRODUCTION

Data analysis is very important in epidemiological research. Currently, there are many commercial statistical software packages widely used by epidemiologists and statisticians around the world. Some charge one time cost and some charge for annual renewal of user's licence. There is also charges for no of user's license. Moreover, frequently new versions of softwares are also launched with additional features and users had to again pay for the new version of software. The continuous expenditure on software is found difficult by most of the researchers and discouraging effect on scientists.

Freely available software packages are limited in number and readiness in use is also difficult. Each one lack in one or other features. EpiInfo, for example, is free and useful for data entry and simple data analysis. Advanced data analysis, however, find it too limited in many aspects. For example, it is not suitable for data manipulation for longitudinal studies. Its regression analysis facilities can not cope with repeated measures and multi level modeling. The graphic facilities are also limited. The 'R' software can handle several datasets simultaneously.

A relatively new and freely available software called '**R**' is promising. Supported by leading statistical experts worldwide, it has almost everything that an epidemiological data analyst needs. It is frequently updated as per advancement in the subject and made available on web free of cost to users worldwide. WHO has found it highly useful and supported number of workshops to train epidemiologists and statisticians in use of '**R**' software in several developing countries.

Recently WHO SEARO office awarded number of fellowships in Epidemiology and Bio-Statistics in India and sent 18 fellows for training in R software at Epidemiology Unit, faculty of Medicine, Prince of Songkla University, Thailand. Participants belonged to various medical colleges and premier medical institutions like ICMR, NIHFWS and NIMS, AIIMS, SGPGIMS etc.

The present workshop jointly organized by the **NIHFWS, NIMS** and is further effort to develop skills among health professionals in the country to use '**R software**' for advance data analysis and use the results to further strengthen the health services and national health programmes. This will be the first workshop on '**R Software**' in India. Depending upon its success, series of other workshops in different parts of the country would be organized.

2. OBJECTIVES

2.1 General Objective

To orient participants in Epidemiological and Bio-Statistical Techniques and develop skills for data management and its analysis in health system and bio-medical research using R Software.

2.2 Specific Objectives

At the end of the course, the participants were able to:

1. Orient participants in basic concepts and methodology in Epidemiology and Bio-Statistics;
2. Orient participants in basics of 'R' Software and its useful features
3. Train participants in database creation and management using 'R' Software
4. Develop skills in participants in advance data analysis using 'R' Software.
5. Train participants for undertaking similar training in different parts of country.

3. WORKSHOP CONTENTS

The workshop was intended to cover topics such as basics of R Software, database management and generation of statistical tables, export and import of data, computation of epidemiological indicators, causation and correlation, confounding and interactions, logistic regression analysis, multivariate regression analysis, modeling techniques, calculation of sample size etc.

4. METHODOLOGY

The methodology of the course comprised of lecture-discussions, demonstration and hands on practice.

5. NUMBER AND NATURE OF PARTICIPANTS

This course was basically meant for faculty members of medical colleges and health institutions in the country . Total 19 participants attended the course. List of participants is enclosed in Annexure 1.

6. WORKSHOP OUTCOME

After the training, the participants were able to understand advance epidemiological and Bio-statistical techniques and its utilization for undertaking health research.

7. DURATION OF THE COURSE: 6 days

8. WORKSHOP FEE

Rs 1000/- were charged from participants to attend the 6 days workshop.

SCHEDULE FOR SIX DAYS WORKSHOP ON 'R' SOFTWARE

Day/Date	Session I	Session II	Session III	Session VI	Session V
Time	9.00-10.00	10.00 -11.00	11.30-1.00	2.00- 3.30	4.00-5.30
Monday May 24, 2010	Registration and Inauguration of the Workshop by Prof Deoki Nandan	Overview of R Software & Graphics (LD & Demonstration)-Prof Virasakdi	Installation of R-Software/Tine-R/Epicalc (LD & Demonstration)-Prof Virasakdi, Dr D Sahu & Team	Basic commands in Epicalc package, Vector and Matrics properties, Reading data file etc (LD & Demonstration)-Prof Virasakdi & Dr D Sahu	Data exploration using Epicalc, Data Management Data cleaning, Missing data, Recode data(LD,Demonstration & Exercise)- Dr D Sahu & Dr B.S.Dhillan
Tuesday May 25, 2010	Concept and Application of Linear and Multivariate Regression (LD)-Prof P.M Kulkarni	Linear and Multivariate Regression using R Software(LD & Demonstration)-Prof Virasakdi	Linear and Multivariate Regression using R Software(LD & Demonstration)-Prof Virasakdi	Confounding and Interaction using R software (LD & Demonstration)-Prof Virasakdi	Modeling Strategy using R Software (LD & Demonstration)-Prof Virasakdi
Wednesday May 26, 2010	Logistic Regression: (LD)-Dr Atul Juneja/Prof V K.Tiwari	Logistic Regression using R Software (LD & Demonstration)-Dr D.Sahu& Dr B.S.Dhillan	Test of Hypothesis (LD)-Prof V.K Tiwari	Test of Hypothesis using R Software(LD & Demonstration)-Dr D Sahu & Dr B.S.Dhillan	Test of Hypothesis using R Software(Exercise)-Dr D Sahu & Dr B.S.Dhillan
Thursday May 27, 2010	Measures in Epidemiology(LD)-Prof DCS Reddy/Prof V K.Tiwari	Measures in Epidemiology using R Software(LD & Demonstration)-Dr B.S.Dhillan	Survival Analysis (LD)-Prof A.Pandey	Life table Kaplan-Meier curve Cumulative hazard rate(LD & Demonstration)-Exercise-Dr D Sahu	Life table Kaplan-Meier curve, Cumulative hazard rate(Exercise)-Exercise-Dr D Sahu
Friday May 28, 2010	Linear and Multivariate Regression using R Software(Exercise)-Dr D Sahu & Team	Confounding and Interaction using R software(Exercise)-Dr D Sahu & Team	Modeling Strategy-Exercise-Dr D Sahu&Team	Modeling Strategy-Exercise-Dr D Sahu&Team	Graphics in R Exercise-Dr D Sahu&Team
Saturday May 29, 2010	Survey Design & Sample size (LD&Exercise) Dr Atul Juneja/Prof V K Tiwari	Sample size, Exercise Prof V.K Tiwari&Team	Clarifications and Feedback on Workshop	Concluding Session : Review of workshop and feedback Distribution of certificates	

RESOURCE PERSONS

EXTERNAL

Prof. Virasakdi Chongsuvivatwong
Epidemiology Unit
Prince of Songkla University
Thailand

Prof. D C S Reddy
TNP-HIV/AIDS(M&E)
WHO Country Office
Nirman Bhavan, New Delhi

Prof. P M Kulkarni
Centre for the Study of Regional Development (CSR D)
Jawahar Lal Nehru University (JNU)
New Delhi-110067

Prof. A Pandey
Director, National Institute of Medical Statistics (NIMS)
ICMR campus, Ansari Nagar,
New Delhi

Dr Atul Juneja
National Institute of Medical Statistics (NIMS)
ICMR Campus, Ansari Nagar,
New Delhi

Dr. D S Sahu
Scientist
National Institute of Medical Statistics (NIMS)
New Delhi

Dr. B.S. Dhillon
Scientist- C
ICMR, New Delhi.

INTERNAL

Prof. Deoki Nandan
Director,
National Institute of Health & Family Welfare
New Delhi – 110067

Prof. V.K. Tiwari
Deptt. Of Planning & Evaluation
National Institute of Health & Family Welfare
New Delhi – 110067

Mr. P.D. Kulkarni
Computer Centre
National Institute of Health & Family Welfare
New Delhi – 110067

Mr. J.C. Sharma
Computer Centre
National Institute of Health & Family Welfare
New Delhi – 110067

LIST OF PARTICIPANTS

1. Dr. Ashfaq Ahmad Bhat
Lecturer Epidemiology
Govt SKIMS Medical College
Bemina, Srinagar, Jammu & Kashmir
Email: drashfaqbhat@yahoo.co.in

2. Dr. Varun Kumar Arora
Senior Resident
Deptt. Of Community Medicine,
PGIMS, Rohtak, Haryana
Email: dr.feats@gmail.com, varun_81in@rediffmail.com

3. Mr. Shushil Kumar Singh
University of Lucknow,
44, Marutipuram, Lucknow,
Uttar Pradesh
Email: shushilksingh@gmail.com

4. Ms Rajkumari Sanatombi Devi
Lecturer in Statistics
Deptt. of Community Medicine,
Sikkim Manipal Institute of Medical Science
5th Mile, Tadong, Gangtok – 737102
Email: rajkumari_sd@yahoo.com

5. Kh. Jiten Kumar Singh
Scientist-B,
National Institute of Medical Statistics
ICMR, Ansari Nagar
New Delhi – 110029
Email: jitensinghk@gmail.com

6. Dr. Raman Chauhan
Research Project Officer (Jr. Resident)
Deptt. of Community Medicine,
IGMC, Shimla – 171001
Himachal Pradesh
Email: drmanchauhan@rediffmail.com

7. Dr. Deepak Sharma
Research Project Officer (Jr. Resident)
Deptt. Of Community Medicine,
Indira Gandhi Medical College,
Shimla – 171001
Email: drdeepaksharma@live.in
8. Dr. Debashis Dutt
Professor, Deptt. of Epidemiology
All India Institute of Hygiene and Public Health
110, C.R. Avenue, Kolkata – 700073
West Bengal
Email: dutta29@gmail.com
9. Dr. Shraddha Mishra
Post Graduate Student
G.R. Medical College,
Gwalior , Madhya Pradesh
Email: drshraddhamishra@yahoo.com
10. Dr. Dipak Pal
Associate Professor
Deptt. Of Epidemiology
All India Institute of Hygiene and Public Health
110, C.R. Avenue, Kolkata – 700073
West Bengal
Email: drdipakpal@yahoo.co.in
11. Dr. Nazish Rasheed
SR (LHMC), Lady Hardinge Medical College
New Delhi
Email: nazish.rasheed@yahoo.co.in
12. Dr. Khan Amir Maroof
Lecturer,
Deptt. Of Community Medicine,
UCMS & GTB Hospital
Delhi.
Email: khanamirmarroof@gmail.com
13. Dr. Ruchi Jain
Senior Resident, LHMC
Deptt. Of Community Medicine
Lady Hardinge Medical College
New Delhi
Email: ruchiafmc@yahoo.com

14. Mr. M.S. Bharambe
Assistant Professor
Deptt. Of Community Medicine
MGIMS, Sevagram, Wardha
Maharashtra
Email: bharambem@gmail.com

15. Dr. M. Hemanta Meitei
Lecturer
Deptt. Of Statistics & Demography
N.I.H.F.W., New Delhi – 110067
Email: mhmeitei@nihfw.org

16. Dr. L.Lam Khan Piang
Deptt. Of Planning & Evaluation
N.I.H.F.W., New Delhi – 110067
Email: lampiang@gmail.com

17. Mr. Sherin Raj T.P.
Deptt. Of Planning & Evaluation
NIHFW, New Delhi – 110067

18. Mr. Subhash Chand
Deptt. Of CHA,
NIHFW, New Delhi – 110067

19. Ms Smita Verma
PT Faculty
NIHFW, New Delhi
Email: smitaverma22@gmail.com

Annexure - 2

Evaluation of the workshop was conducted on the last day of the training programme. A series of questions regarding conduction of the training, training arrangements, lodging and boarding arrangements. etc. were asked from participants. Each participant was given a questionnaire to give his/her feedback.

Following findings emerged after analysis of the feedback:

1. All the trainees agreed that course objectives were fulfilled and agreed that methodology adopted during the course was appropriate, but some did not agree, as they wanted more time for practical exercises, which was currently not possible. All the trainees agreed that course contents were fully covered.
2. All participants agreed that the course would be of great use to them. Some expressed that Syntax writing require more time and practice, as they did not know this before.
3. All participants find sessions covered in the workshop useful and no topic was asked for deletion. All participants appreciated the course as no sessions were left in the training.
4. Most of the participants found the hostel, computer lab and food arrangements acceptable. The participants rated the computer lab facility as excellent. They appreciated the effort made by faculty and staff of computer center for the smooth conduction of the training programme. Overall arrangements were also rated as very good.