

Applied Epidemiology Using R

*UC Berkeley School of Public Health
Syllabus for PH 298 (36) CC# 76592
Fall 2007, Modified 2007-08-27*

Faculty

Tomas Aragon, MD, DrPH, Executive Director, Medical Epidemiologist
Center for Infectious Disease Preparedness, UC Berkeley School of Public Health
Email: aragon@berkeley.edu, Tel: 510-643-4935; Fax: 510-643-4926

Wayne Enanoria, PhD, MPH, Director, Epidemiology Preparedness & Informatics
Center for Infectious Disease Preparedness, UC Berkeley School of Public Health
Email: enanoria@berkeley.edu, Tel: 510-643-4934; Fax: 510-643-4926

Michael Samuel, DrPH, Chief, Surveillance and Epidemiology
STD Control Branch, California Department of Health Services
Email: msamuel@dhs.ca.gov,

Schedule

Date	Wk	Topics	Lecturer
Aug 28	1	Getting started with R	Tomas Aragon
Sep 04	2	Working with R data objects I (Vectors, Matrices, Arrays)	Tomas Aragon
Sep 11	3	Working with R data objects II (Lists, Data Frames)	Wayne Enanoria
Sep 18	4	Managing epidemiologic data I (Entering, Editing, Transforming, Merging, Exporting)	Tomas Aragon
Sep 25	5	Managing epidemiologic data II (Importing, Dates, Missing values)	Tomas Aragon
Oct 02	6	Analyzing simple epidemiologic data	Tomas Aragon
Oct 09	7	Descriptive epidemiologic methods	Tomas Aragon
Oct 16	8	Analytic epidemiologic methods (includes multivariable regression)	Tomas Aragon
Oct 23	9	Graphing epidemiologic data I	Michael Samuel
Oct 30	10	Graphing epidemiologic data II	Michael Samuel
Nov 06	11	R programming I	Tomas Aragon
Nov 13	12	R programming II	Tomas Aragon
Nov 20	13	Sample size calculations	Wayne Enanoria
Nov 27		Thanksgiving	
Dec 04	14	Complex Survey Analysis using the Survey package	Wayne Enanoria
Dec 11	15	Student presentations	Tomas Aragon

Course description

This is an intensive one-semester introduction to the R programming language for applied epidemiology. R is a freely available, multi-platform (Mac OS, Linux, and Windows, etc.), versatile, and powerful program for statistical computing and graphics (<http://www.r-project.org>). This course will focus on core basics of organizing, managing, and manipulating epidemiologic data; basic epidemiologic applications; introduction to R programming; and basic R graphics.

Target audience

This course is intended for epidemiologists, medical epidemiologists, data analysts, and demographers that want an introduction to the R language for epidemiologic applications.

Course prerequisites

Completion of one semester of epidemiology and one semester of bio/statistics.

Course objectives

Upon completion of this course, participants will be able to:

- Use R as a scientific calculator and a functional spreadsheet;
- Enter, manage, and manipulate epidemiologic data in R;
- Conduct basic epidemiologic analyses in R.;
- Graphically display epidemiologic data;
- Write basic R programs.

Course format

Lecture and computer demonstration. You are welcome to bring your laptop with Rpre-installed.

Course enrollment and fee

UC Berkeley students should register for Public Health 298, Sect. 36, CCN 76592. Non-registered students who want to receive academic credit will need to register and pay the UC Extension fee (see <http://www.unex.berkeley.edu/info/concur.php>). For all others: to enroll in this course, follow instructions posted on our website at <http://www.idready.org>, or show up to the first day of class.

CIDP follows the UC Berkeley academic calendar. For the complete Academic Calendar, go to <http://opa.berkeley.edu/AcademicCalendar/calendar.cfm>.

Course credit/Grading

For registered UC Berkeley/Extension students: Units: 2; Grading: P/NP

Course location and schedule

Schedule: Every Tuesday, 1:00 pm - 3:00 pm, starting August 28, 2006.

Location: 1918 University Avenue, 4th Floor (2.5 blocks west of campus)

Course books

- Aragon TJ, et al. Applied Epidemiology Using R. Available at <http://www.medepi.net/epiwithr/>
- Myatt Mark. Open Solutions – R. Nordic School of Public Health. June 2004 (Available for free at <http://www.brixtonhealth.com/Rex.zip>)
- An Introduction to R. Freely available at <http://cran.r-project.org/doc/manuals/R-intro.pdf> (also comes with default installation)

Course requirements and evaluation

- Class attendance and participation
- Complete and present student project (see below)

Student project

A student project will require the following:

- Read in data set
- Data cleaning, processing, manipulation, transformation, recoding, discretization, etc.
- Basic descriptive method
- Analytic methods (including using regression methods)
- Basic graphical display
- Write and use your own functions

Additional readings & resources:

- Official R manuals at <http://cran.r-project.org/manuals.html>.
- Contributed R tutorials at <http://cran.r-project.org/other-docs.html>.
- Selvin S. Modern Applied Biostatistical Methods: Using S-Plus. Oxford Press; 1st edition (1998) ISBN: 0195120256

Related spring courses

The following are excellent, highly recommended courses:

- PH 248, Steve Selvin, Statistical/Computer Analysis Using R, Tu 10a-12p, 2304 Tolman, CCN: 76178, Units/Credit: 3
- PH 252D, Kathleen Mortimer, Introduction to Marginal Structural Models, MW 10a-12p, 2301 Tolman, CCN: 76216 & 76225, Units/Credit: 4

R for epidemiologic computing

Where do I get R?

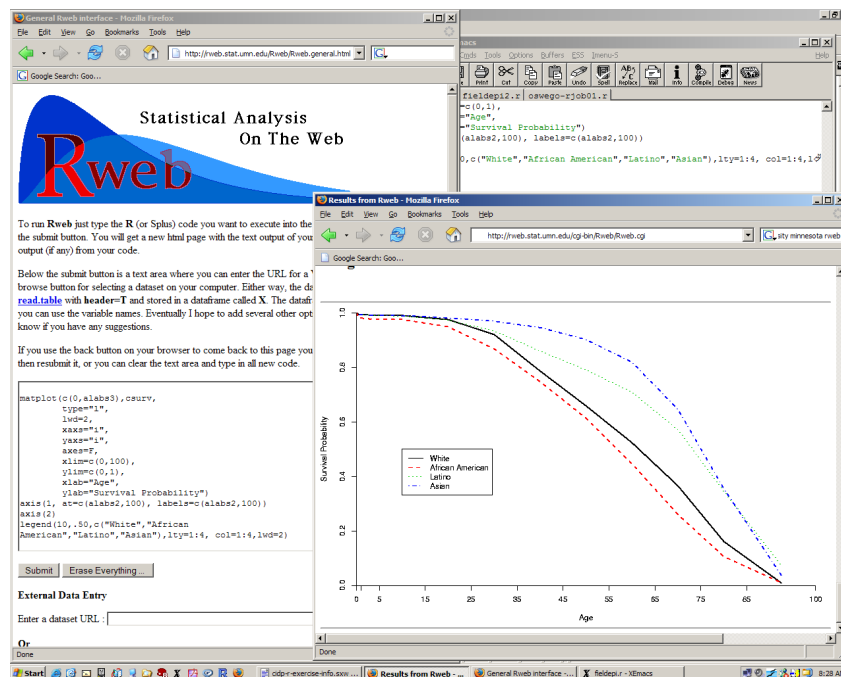
- R project home page at <http://www.r-project.org>
- Download from UC Berkeley at <http://cran.cnr.berkeley.edu/>
- R manuals available at <http://cran.r-project.org/manuals.html>
- R tutorials available at <http://cran.r-project.org/other-docs.html>

How can I get help with R?

- Visit <http://www.epitools.net> (site developed and maintained by Tomás Aragón)
- Join the EpiTools Yahoo Group at <http://www.epitools.net>
- Join an R mailing list at <http://www.r-project.org/mail.html>
- Get great reference card at <http://www.rpad.org/Rpad/R-refcard.pdf>

How can I access R on-the-road?

From any Internet-connected computer you can access Rweb via the Montana State University website at <http://www.math.montana.edu/Rweb/The 'epitools' R package>



EpiTools R Package

Try the 'epitools' R package can be installed from within R. This will be covered in class.