

APPLIED EPIDEMIOLOGIC DATA ANALYSIS
Summer Quarter 2006

Time/Date: Mondays and Wednesdays, 6 – 8pm
July 5 – Aug 23, 2006

Location: 680 N. Lake Shore Drive, Suite 1220 conference room

Course instructor: Mercedes Carnethon, Ph.D.
Assistant Professor of Preventive Medicine
Feinberg School of Medicine
Northwestern University

Office: 680 N Lake Shore Drive, Suite 1102
Office Hours: Wednesdays 12 – 5 and by appointment
Phone: 503-4479
carnethon@northwestern.edu

Course Description

This course is designed to provide hands-on experience in the design and analysis of epidemiologic studies, with emphasis on applying epidemiology and biostatistics methodology to address current clinical and population science questions of interest. Through independent projects and class exercises, students will learn the practical aspects of conducting research, ranging from defining a statistically testable hypothesis to building a database and generating results. A brief review of biostatistical methods will be provided in the context of appropriate statistical tests to address given statistical questions; however, the focus of the course is not on biostatistical methodology as students are expected to have adequate preparation.

The centerpiece of the course is the individual project. MPH students are encouraged to bring ideas for their Culminating Experience to the course so that they can further develop and refine their research plan. MSCI and postdoctoral fellows should bring ideas for studies that they would like to conduct. Students are also invited to develop a “hypothetical” or future study, as the emphasis of the course is in developing the skills necessary to conduct an independent research project in a range of settings (e.g., academia, industry, public health agencies). During this process, students will be guided in the use of current technology to build databases, questionnaires, and conduct data analyses. The target audience for this course is advanced MPH and MSCI students and postdoctoral fellows who plan to conduct independent research.

Course Objectives

1. Define a study question in response to an identified need in the current literature and develop a statistically testable hypothesis
2. Develop a data collection form/questionnaire to gather information required to answer the study question
3. Develop an analytic plan to describe the study population and test at least 3 statistical hypotheses of interest in the study
4. Build a database in Access to collect the data in preparation for analyses

Using a dataset that I will distribute, students will:

5. Prepare a dataset for analysis by “cleaning” and “checking” data for errors
6. Conduct statistical analyses and prepare written responses to a series of questions that I will distribute to the class

The objectives of this course will be met through a combination of exercises that the student performs on a study question of his/her choosing and through a dataset that I will distribute to the class. Ideally, students will not only have developed the skills to conduct independent research, but will have a tangible analytic approach for conducting a research project or proposing a new research study.

Teaching Format

In-class instruction will include a combination of lectures to introduce concepts and exercises to stimulate discussion.

Student Evaluation

Individual Project (60%):

- Regular take-home assignments in the first two-thirds of the course will focus on learning how to complete Objectives 1 through 4. Each objective will be divided into smaller, focused segments and students will be asked to demonstrate their progress on these intermediate steps through assignments to be discussed in class or handed in.
- In addition to homework exercises to practice these concepts, students will be required to demonstrate mastery of these concepts by generating researched, well-written products

Database development and analysis (30%):

- To meet Objectives 5 and 6, students will complete assignments based on the dataset I will distribute (i.e., class project). Because this is not a statistical computing course, example coding will be provided for two popular data analysis systems, SAS and STATA, though other appropriate (Microsoft Excel may NOT be used for data analysis in this course) packages may be used (i.e., SPSS).
- Clarity of written expression will be evaluated

Attendance/Participation (10%):

- Because in-class exercises and discussions will be a central opportunity for students to explore research challenges and introduce new concepts, attendance is crucial. Students who miss more than 2 sessions will be penalized.

There will be no exams in this course.

Course Evaluation

The MPH Program administers web-based course evaluations to students for each course near the end of the quarter. ***Your completion of both the unit (course) and faculty evaluation components is required; failure to complete either of the evaluations will result in an incomplete grade until the evaluations are submitted.*** You will be sent the web link and instructions via email later in the quarter. You will have about two weeks time to complete the evaluations before grades are submitted.

Textbooks

Required:

1. Hulley SB, Cummings SR, Browner WS, Grady D, Hearst N, Newman TB. Designing Clinical Research. 2nd Edition. Philadelphia, PA: Lippincott Williams & Wilkins, 2001.

Recommended:

1. Cody RP, Smith JK. Applied Statistics and the SAS Programming Language. 4th Edition. New Jersey: Pearson Education, 1997
2. Rabe-Hesketh S, Everitt B. A Handbook of Analyses Using STATA. 3rd edition. Chapman & Hall, 2004.

3. Cronan J, Anderson V, Anderson BB. Microsoft Office Access 2003 QuickSteps. McGraw-Hill Osborne Media, 2004.
4. Browner, WS. Publishing and Presenting Clinical Research. Philadelphia, Pa, Lippincott Williams & Wilkins, 1999.

Although not required, access to an individual laptop with a wireless connection will facilitate completion of in-class exercises. In addition to the required textbook, readings will occasionally be distributed in class.

Course Sessions

Date	Topic	Class Activity	Readings*	Assignments Due
7/5	Course overview and study designs	Introductions; Pairing by subject area for later class projects; Class activity to identify questions based on published research to be distributed in class	Hulley Chapters 1-2 Chapters 7 – 13 (as applicable)	
7/10	Hypothesis development	Discuss examples of good/bad research questions; Brainstorm research hypotheses in pairs		
7/12	Measurement assessment 1: objectives, defining exposure and outcome	Discuss research hypotheses and goals turned in;	Hulley Chapter 4	Hypothesis statement with 2-3 supporting paragraphs stating the importance of this research question in the context of current literature
7/17	Measurement assessment 2: reducing error and bias, designing questionnaires	Questionnaire development activity	Hulley Chapter 15	
7/19	Database development 1: data coding, dictionaries	Building a relational database;	Hulley Chapter 16 and other Handouts	List of all variables needed for individual project and type of variable (i.e., continuous/ categorical)
7/24	Database development 2: using Microsoft Access	Class examples and exercises on how to use Access to make forms; Discuss data tables for individual projects	Handouts	
7/26	Computer lab session: form development for individual projects	Class time to build questionnaires and data collection forms	Handouts	Listing of the data tables needed for the database
7/31	Statistical analysis review	Statistics examples; Discussion and work on analyses plans for the individual projects (in pairs) Model building goals and examples;	Chapter 17 Class handouts	
8/2	Analytic decision making: variables and model building	Examples from published research; Develop analysis plan for distributed dataset;	Handouts	
8/7	Statistical power/ sample size	Sample size calculation goals; Sample size examples using analysis packages	Hulley Chapters 5, 6	Tested and formatted database w/ data dictionary
8/9	Presentation of individual projects	10-minute presentations by each individual on proposed project		
8/14	Overview of new dataset (Class Project)	As a group, select 3 hypotheses of interest in the new dataset		

Date	Topic	Class Activity	Readings*	Assignments Due
8/16	Data cleaning and checking	Review the distribution of the variables of interest and generate plots; Begin identifying the distribution of the variables in the distributed dataset;		Data analysis plan and power/ sample size analysis due for individual project
8/21	Computer lab session to complete project analyses		Handouts	
8/23	How to interpret findings		Handouts	Analysis results from class project due
8/30	No Class—final assignments due			1. Written interpretation of findings from class project 2. 350 word abstract of findings from class project

*Readings are to be completed in advance of the course session that they correspond to.