PUB HLTH 437  
PRACTICUM ON EPIDEMIOLOGIC RESEARCH DESIGN AND DATA ANALYSIS – 1.0 CREDIT  
Summer Quarter 2014

Time/Date:  This class will be taught as an independent study. We will meet on the first scheduled day of classes, **Monday June 23**, to discuss our schedules and how assignments will be completed. Class is scheduled to meet between 6-9 pm.

Location:  McGaw 1-401 (June 24)  
Galter Library LRC Lab (Dates TBA)

Course instructor:  Suzanne E. Belinson, MPH, PhD  
Adjunct Assistant Professor of Preventive Medicine  
Feinberg School of Medicine  
Northwestern University

Office:  225 North Michigan Ave  
Office Hours:  By appointment  
Office phone:  312-297-5559 (voice mail)  
Email:  suzanne.belinson@bcbsa.com

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 generating results from applied data analyses. The overarching goal of this course is for students to develop the basic skills and to have a tangible analytic approach to conduct independent empirical research involving statistical data analysis. 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 MUST bring ideas for their Culminating Experience to the course so that they can further develop and refine their research plan. MSCI, postdoctoral fellows, and all other students MUST bring ideas for studies that they would like to conduct. 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 refine research hypotheses and to conduct data analyses. The target audience for this course is advanced MPH and MSCI students and postdoctoral fellows who plan to conduct independent research.

Prerequisite

Intro to Biostatistics and Intro to Epidemiology

Course Objectives

1. Define a study question in response to an identified need in the current literature and develop a statistically testable hypothesis
2. Choose suitable study populations and measurements for the research question under study
3. Demonstrate some basic understanding of questionnaire formation, data collection process, and database development

4. Develop an analytic plan to describe the study population, outcomes, and test the statistical hypotheses of interest in the study

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 preparing written responses to a series of questions that I will distribute to the class.

Teaching Format

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 (“individual project”) and through a dataset that I will distribute to the class (“class project,” with individual choices of study hypotheses and variables within the framework of the dataset). If you have a dataset for your CE you can use that dataset to complete the final project.

You will be given a set of lecture notes covering various topics to introduce concepts and practical examples and exercises to stimulate discussion. We will schedule to meet a few times throughout the term for class discussion. Class discussion will focus on a topic central to the completion of the individual project. Students are strongly encouraged to actively participate in these discussions in order to: 1) improve oral communication skills; 2) revise and strengthen their individual projects through feedbacks from the instructor and their peers; 3) benefit from an exposure and contribution to other topics their peers are conducting. The sessions on the class project will be conducted in the computer lab. Step-by-step instructions using STATA will be provided in the lab to facilitate completion of the class project.

Student Evaluation

Individual Project (40%):
- Regular take-home assignments 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 handed in and discussed in class.
- 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. Two of which will be graded on (1-page research proposal – 10% and data analysis plan - 20%).
- Each student will be asked to give an oral presentation with PPT on their individual project. This presentation -- about 5-10 minutes (followed by a short discussion) -- will account for 10% of the total score.

Class Project (40%):
- 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 STATA only – the required software for the MPH program, though other appropriate packages may be used (e.g., SAS, SPSS). Microsoft Excel may NOT be used for data analysis in this course. As stated above if you have a dataset for your CE you may use it for the class assignment. This must be cleared with me prior to beginning the analysis.

For both individual and class projects, clarity of written expression will be evaluated. Due dates are clearly specified in this syllabus (see course sessions table on pages 4 and 5). Late assignments will be penalized: up to 1 week late, 10% reduction; up to 2 weeks late, 20% reduction; longer than 2 weeks, 30%.

Attendance and Participation (20%):
- Because in-class exercises and discussions will be a central opportunity for students to explore research challenges and introduce new concepts, attendance and active participation are crucial. Attendance and
participation will each account for 10% of the grade. Attendance will be 0 for those missing more than
2 sessions. This will only apply to the sessions that we schedule on the first day of class. This class is
independent study.

There will be no exams in this course. The University has a deadline of Aug. 28th to provide grades for all
summer courses.

**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.

**Academic Integrity**
Every Northwestern faculty member and student belongs to a community of scholars where academic integrity is
a fundamental commitment. The Program in Public Health abides by the standards of academic conduct,
procedures, and sanctions as set forth by The Graduate School at Northwestern University. Students and faculty
are responsible for knowledge of the information provided by The Graduate School on their Web page at
http://www.tgs.northwestern.edu/academics/academic-services/integrity/index.html

Academic misconduct includes, but is not limited to
1. Receiving or giving unauthorized aid on examinations or homework
2. Plagiarism
3. Fabrication
4. Falsification or manipulation of academic records
5. Aiding or abetting any of the above

The PPH follows The Graduate School’s procedure for evaluating alleged academic misconduct, as outlined on
the TGS website. http://www.tgs.northwestern.edu/academics/academic-services/integrity/dishonesty/index.html

Faculty reserve the right to use the “Safe Assignment: Plagiarism Detection Tool” that is part of the Course
Management System to evaluate student assignments. Information about this tool can be found at
http://www.it.northwestern.edu/education/course-management/support/assessments/safeassignment.html

**Textbooks**

There are no required textbooks for this course. The following are recommended resources.

**Recommended:**


   2006.

   Williams & Wilkins, 2006.

In addition to the required textbook, readings will occasionally be distributed in class.
Communications

All registered students have been added to the course management system (“blackboard”) (https://www.courses.northwestern.edu/webapps/login). The main means of communication will be via email. I can be reached at Suzanne.belson@bcbsa.com

Course Sessions

See next page. Readings should be completed according to this schedule to keep on target.

There will be 12 modules. The first will meet at McGaw 1-401. The last 3 (in August) labeled with “computer lab session” on the next page) will meet at the Galter library Learning Resources Center lab (“teaching space”). It is located on the 1st floor of the Galter Library in the Ward building. Ask the front desk if needed. We will discuss these dates. This is written as if we were meeting in class. On the first day we will discuss some options.
<table>
<thead>
<tr>
<th>Date</th>
<th>Topics</th>
<th>Class Activities and Handouts</th>
<th>Readings or Handouts</th>
<th>Assignments Due (and Notes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6/23 Mon.</td>
<td>Course overview; Introduction to research question, study design, study population</td>
<td>Introduction; Course Overview; Review of study designs; Criteria for defining a study question; <strong>Handouts</strong>: class discussion assignment 1; CE excerpt (online)</td>
<td>Hulley Chapters 1-2; Chapters 7 – 12 (as applicable)</td>
<td>In-class mini-survey of research experiences, interests, and course expectations</td>
</tr>
<tr>
<td>6/25 Wed.</td>
<td>Hypothesis development; Measurement assessment</td>
<td>Discuss research questions and design; Lecture on hypothesis development; Defining the exposure and outcome; Reducing error and bias; <strong>Handout</strong>: class discussion assignment 2</td>
<td>Hulley Chapters 3-4</td>
<td>In-class discussion of your research question, study design and population</td>
</tr>
<tr>
<td>6/30 Mon.</td>
<td>From Design to Data – the Long Journey in Between; Human Subjects/Institutional Review Board</td>
<td>Discussion of hypotheses and measurements; Brief overview of manual of operations, data collection and database management; Human Subjects/IRB discussion; <strong>Handouts</strong>: 1-page written proposal instructions; Rules for better data collection (online)</td>
<td>Hulley Chapter 14-17;</td>
<td>In-class discussion of research hypotheses and measurements, and other practical issues in carrying out your research project</td>
</tr>
<tr>
<td>7/2 Wed.</td>
<td>Statistical analysis review; Generating an analysis plan; Sample size calculations</td>
<td>Brief review of applied statistical analysis; Example of an analysis plan; Overview of sample size calculations; <strong>Handout</strong>: class discussion assignment 3; stat. test guide</td>
<td>Hulley Chapters 5-6;</td>
<td>1-page written draft describing the proposed study DUE (Graded)</td>
</tr>
<tr>
<td>7/7 Mon.</td>
<td>Secondary data analysis (research using existing data)</td>
<td>Discuss statistical analysis plans; secondary data analysis; introduction to available data sources at NUPM and beyond; <strong>Handout</strong>: data analysis plan example (scoring guide)</td>
<td>Hulley Chapter 13</td>
<td>In-class discussion of independent project statistical analysis plans.</td>
</tr>
<tr>
<td>Date</td>
<td>Topics</td>
<td>Class Activities and Handouts</td>
<td>Readings or Handouts</td>
<td>Assignments Due (and Notes)</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>7/9 Wed.</td>
<td>How to read, review apply the medical literature</td>
<td>Discuss how to examine the literature. examples in the published literature and the</td>
<td>Read the papers on Blackboard</td>
<td>In-class discussion of these papers.</td>
</tr>
<tr>
<td>7/14 Mon</td>
<td>Introduction to Microsoft ACCESS;</td>
<td>ACCESS as database management tool</td>
<td>Guest Speaker: Daniel B. Garside</td>
<td>Data Analysis Plan DUE (graded);</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Handout:</strong> presentation scoring guide; proposal assignment instruction;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7/16 Wed.</td>
<td>The Fundamentals of Questionnaire Design</td>
<td>Lecture and in-class exercises on questionnaire design</td>
<td>Guest Speaker(s): Elizabeth Hahn</td>
<td>Bring your own questionnaire or issues related to questionnaire design to class (optional)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBA</td>
<td>Presentation of individual projects; Introduction/overview of new dataset</td>
<td>5 to 10-minute presentations by each individual on proposed project with time for Q&amp;A and feedback; <strong>Handout:</strong> class data analysis project example</td>
<td></td>
<td>Individual presentation with PPT (graded)</td>
</tr>
<tr>
<td>TBA</td>
<td>Computer lab session: Outcome, exposure, and covariates measurement, preliminary analyses and Examining basic associations;</td>
<td>Data cleaning and checking; Review the distribution of variables of interest and generate plots; Conduct preliminary analyses.</td>
<td>Handout on STATA</td>
<td>A study hypothesis based on the dataset should have been formulated before class: Be ready to start the initial data analysis</td>
</tr>
<tr>
<td>TBA</td>
<td>Computer lab session: Confounding and interactions; Multivariable model building</td>
<td>Hands-on session on using STATA to examine ways to deal with confounding and interaction; Conduct multivariable analyses.</td>
<td>In-lab completion of relevant sections of the class project</td>
<td>Analysis and written interpretation of findings from class project, Part I DUE (graded)</td>
</tr>
<tr>
<td>TBA</td>
<td>Computer lab session to complete project analyses final assignments due at 9pm</td>
<td>Supervised data analysis lab session to complete analysis for project; Late assignments will be penalized</td>
<td>Lab reserved for this course</td>
<td>Required Course evaluation DUE (without which no grade for the course can be given) Analysis and written interpretation of findings from class project, Part II DUE (graded)</td>
</tr>
</tbody>
</table>