EPID 766: Epidemiologic Research Using Healthcare Databases

COURSE INFORMATION

Class Time: Tuesdays/Thursdays 3:30-4:45pm
Location: McGavran-Greenberg Hall, Room 1304

Lead Instructors

Jennifer Lund
Department of Epidemiology, UNC School of Public Health
2102D McGavran-Greenberg
Phone: (919) 966-7440 | Fax: (919) 966-2089
Email: Jennifer.Lund@unc.edu
Office hours: By appointment

Alan Kinlaw
Division of Pharmaceutical Outcomes and Policy, UNC School of Pharmacy
2201 Kerr Hall
Phone: (919) 966-2747 | Fax: (919) 966-8486
Email: akinlaw@unc.edu
Office hours: By appointment

Teaching Assistants

Arti Virkud: avirkud@live.unc.edu
Paula Strassle: pstrass@live.unc.edu
Office hours: Tuesdays from 1:00-2:00pm (Arti): McGavran-Greenberg 2305
Wednesdays from 2:00-3:00pm (Paula): McGavran-Greenberg 2303
Thursdays from 12:15-1:15pm (Arti): Rosenau 231 (except Feb 7 in Ros 235)
Thursdays from 1:45-2:45pm (Paula): McGavran-Greenberg 2305

COURSE DESCRIPTION AND PREREQUISITES

Prerequisite, competency in data management with SAS (e.g., EPID 700, BIOS 511, or equivalent) and knowledge of basic epidemiology (e.g., EPID 600 or 710). The course will teach students how to use administrative healthcare utilization data for epidemiologic studies. Students will gain an understanding of how the data are generated through claims for reimbursement sent to payors from physicians, hospitals, and pharmacies. The common elements found in administrative healthcare utilization data will be described, including outpatient medical insurance claims, hospitalization information, and pharmacy claims. Students will learn about coding nomenclature, including ICD-9 CM diagnosis codes, CPT procedure codes, and various medication coding systems. Students will be given access to a sample of healthcare claims data and will learn how to use the data to identify populations of interest and conduct epidemiologic studies of the utilization and comparative effectiveness/safety of prescription drugs and healthcare services. Students will also learn about current hot topics in healthcare database research including the use of electronic health record and international databases, the transition to ICD-10 coding schema, common data models and distributed data networks, and software tools for data visualization.

DESIRED COURSE OUTCOMES

- To understand the contents and structure of typical healthcare insurance claims databases
- To demonstrate facility with information technology processes to ensure secure use of data resources
- To efficiently manipulate healthcare claims data using SAS to construct analytic cohorts
- To implement a basic descriptive epidemiologic study using healthcare claims data

JOURNAL CLUB DISCUSSION AND LAB ASSIGNMENTS

Journal club discussions
Throughout the semester, students will take part in 7 journal club discussions focused on published research studies using the Truven Health Analytics MarketScan data. These articles have been selected to highlight specific issues important for database research. All students are expected to read the assigned articles and actively participate in discussion. See course schedule and Sakai site for further details.

Lab assignments
Students will also be responsible for completing 6 lab exercises. These assignments focus on developing a necessary aspect of an epidemiologic cohort study. Students should upload the SAS program and Word document to the Sakai site under the "Assignments" section on the specified due date by 3:25pm. All SAS programs should include a header (see file header.sas on Sakai for appropriate format) and should be richly commented so that the intent of the program is obvious to the instructors and TAs.

PROJECT

As part of EPID 766, each student will conduct a descriptive epidemiologic study using the Truven Health Analytics 1% sample of the MarketScan Commercial Claims and Encounters Database. Ideally, the course project will be related to the student's research interests. **Working in pairs or groups of three is encouraged.** For students without a developed research interest, the course instructor can help to suggest a topic.

There will be 3 **homework** assignments to help students prepare their final research proposal for the project. After incorporating course instructor, TA, and peer review feedback, student(s) will upload their final 1-2 page (single spaced) research proposal to the Sakai site under the "Assignments" section on the specified due date by 3:25pm. This project proposal should contain the following sections: 1) Background, 2) Specific Aims, and 3) Methods and include references. Sample project proposals are provided on Sakai.

The results of this research project should be submitted as a report on the final day of the class. The report should be similar to a brief scientific communication and should contain the following sections: 1) Introduction; 2) Methods; 3) Results; 4) Discussion/Conclusions. However, given that this not a fully developed research project, the introduction and discussion sections should be short (1-2 paragraphs each). The entire report should not exceed 2,500 words. The SAS code used

to generate the results should be submitted separately to the Sakai site and will also be graded. As with the lab assignments, make sure to use good SAS commenting.

Finally, students will present results of their final project in one of the last 3 days of the course. The presentations should be similar to those at the annual meeting of the International Society for Pharmacoepidemiology. Students will be allotted a 15 minutes time slot. They should plan to present for 10 minutes and reserve the final 5 minutes for questions. Students who work in pairs will be expected to undertake projects that are somewhat more substantial than projects done by individual students.

The project grade will be assigned based on the full research proposal (15%), final project presentation (25%), final project report (35%) and accuracy/clarity of the SAS program (25%).

GRADING

The final course grade will be based on journal club participation (15%), homework assignments related to project proposal (15%), labs (20%) and the final project and presentation (50%).

TOPICS TO BE ADDRESSED

- Introduction to epidemiologic research using health care data
- Orientation to secure access to the pharmacoepi server environment
- US Healthcare system, types of healthcare data, and payment systems
- Medical coding systems (diagnoses, procedures, and prescription drugs)
- Introduction to Proc SQL in SAS
- Resources useful in developing claims-based definitions of health exposures and outcomes
- Using outpatient and hospitalization claims to identify incidence and prevalence of diseases
- Using prescription drug claims to identify new users and patterns of drug use
- Using claims to identify comorbidities and other measures of physical function
- Estimation of costs of care
- Estimation medication adherence and persistence

DATA PRIVACY AND DATA USE AGREEMENTS

The course is taught using proprietary data from Truven Health Analytics. Although direct patient identifiers have been removed from the data, the data contain dates of services and therefore it may be possible to identify individual patients. **For this reason, the data must remain on the server and cannot be copied onto individual computers FOR ANY REASON.** Prior to being granted access to the data, students will be required to sign a contract that governs use of the data and our computing system. Deliberate violation of the terms of this agreement would have serious consequences for everyone at UNC using the Truven Health Analytics data for research (immediate termination of our data use agreement) and would expose you to potential civil and/or criminal penalties. Please talk to the course instructors or TAs if you have any questions about appropriate use of data.

COMPUTING SYSTEM

Students will be given temporary accounts that will allow them to use the research computing environment established by the Pharmacoepidemiology program area in the Department of Epidemiology and administered by the Sheps Center for Health Services Research. The students will be sharing this resource with programmers, faculty, and other graduate students using this computer for dissertation research. Students are requested to use the system only for class work. Student accounts will be closed at the end of the semester and all files saved on the computer will be deleted. If students want to keep copies of programs, they should download these to a USB drive before the last day of the semester. **Again, data cannot be taken off the system**. All use of the system is logged.

It is possible to gain access to the data after the end of the course for research purposes. Please talk to the course instructors to learn more about data access.

FINAL NOTE ABOUT COURSE

This course can be technologically challenging at first. We have worked hard to improve the interface to our secure computing environment that will permit all of the students in the class to simultaneously access the database. However, it is possible that there will be technological/computing problems that will occur during the course of the semester. If problems arise, for example, if you have problems accessing data or running programs, please let the course instructor and/or TAs know as soon as possible. We will work to resolve the problems or develop work-around solutions. In the event of problems with the data or computing system that cannot be fixed, we will modify the expectations and/or due dates of impacted assignments.

IMPORTANT NOTE ABOUT FACULTY RESOURCES

Alan Kinlaw is a designated Responsible Employee at UNC-Chapel Hill. This kind of designation is sometimes referred to as a "mandatory reporter," but our university uses the term "Responsible Employee." When someone discloses an incident of discrimination, harassment, or retaliation based on any protected status, sexual assault or sexual violence, sexual exploitation, interpersonal violence, or stalking to a Responsible Employee, that employee must report that information to any professional staff member in UNC's Equal Opportunity and Compliance Office. The Responsible Employee must share as much information as they know, such as the date, time, and location of the incident, the names of the parties involved; a brief description of the incident; and whether the incident has already been reported to someone else.

The purpose of reporting an incident is to help ensure the safety and well-being of the person(s) affected. When EOC receives a report, either a Report and Response Coordinator, the Title IX Compliance Coordinator, or an EOC staff member will reach out to the affected person via email to let them know about their rights to seek medical care and notify law enforcement, and their options for seeking support and reporting the incident to the University. If the report involves a student as the affected person, the Report and Response Coordinator will reach out to the student. If the report involves an employee as the affected person, another staff member from

EOC will contact the employee. The affected person can choose whether they want to meet with someone from EOC to discuss interim protective measures (e.g., no contact orders, changes to housing, changes to work schedules, academic accommodations) and reporting options, but they are not required to do so.

For more resources and information about discrimination, harassment, sexual violence, interpersonal violence, and stalking, you may want to visit UNC's designated website: https://safe.unc.edu/.