COLUMBIA | MAILMAN SCHOOL OF PUBLIC HEALTH

P8400 Epidemiology III: Applied Epidemiologic Analysis
Fall Semester 2019

COURSE DESCRIPTION

This is the third epidemiology course in the sequence of methods courses. The course focuses on integrating study design methods with advanced statistical analyses. Lectures cover theoretical concepts, including confounding, interaction, and pseudo risks and rates, and several analytical methods including linear, logistic, Cox Proportional Hazards, and Poisson regression methods. The course will also touch briefly on other generalized linear models (e.g. relative risk regression) and methods for handling correlated data (GEE, multi-level modeling). Computer laboratories make use of multiple datasets to provide students with an opportunity to implement the analytical methods covered in lecture. Multivariable methods to test for confounding, interaction, and mediation are taught both in lecture and laboratories. The goal of this course is to provide both theoretical and practical experience in analyzing and interpreting epidemiological data.

COURSE LEARNING OBJECTIVES

Students who successfully complete this course will be able to:

1) Compare and differentiate the strengths and limitations of the different study designs in epidemiology
2) Examine and quantify the role of information bias, selection bias, and confounding in the estimation of measures of association
3) Evaluate and quantify the role of interaction on both an additive and multiplicative scale and the implications of scale on estimation
4) Apply multivariable regression methods commonly used in epidemiology and understand the assumptions of these regression methods compared to categorical methods
5) Assess confounding, interaction and mediation using multivariable models

PREREQUISITES

Epidemiology II
Analysis of Categorical Data
Applications of Epidemiologic Research Methods

CLASS SESSIONS

Tuesdays: 1:00 PM – 3:50 PM

- Lecture: 1:00 PM – 2:50 PM, Hammer Health Sciences Center (HSC), Room 301
- Lab: 3:00 PM – 3:50 PM, Hammer Health Sciences Center (HSC)
  - Lab Section 001: HSC LL106
  - Lab Section 002: HSC LL107
  - Lab Section 003: HSC LL108A/B
  - Lab Section 004: HSC LL109A/B
  - Lab Section 005: HSC LL110

INSTRUCTOR

Jeanine Genkinger, PhD
Phone: 212-342-0410
Email: jg3081@cumc.columbia.edu
Office: 722 West 168th St, 7th Floor, Rm 712
Office Hour: By Appointment

TEACHING ASSISTANTS

Teaching Assistant	Email	Office Hour Time	Office Hour Location
Adiba Ashrafi	ai2337@cumc.columbia.edu	TBD	TBD
Precious Esie	pie2104@cumc.columbia.edu	Tuesdays, 11 AM – 12 PM	ARB 15th floor conference room (1514)
Diana Garofalo	dcg2132@cumc.columbia.edu	Mondays, 12 PM – 1 PM	ARB 7th floor conference room (739A)
Shadiya Moss	sm4248@cumc.columbia.edu	Mondays, 11 AM – 12 PM	ARB 7th floor conference room (739A)
John Pamplin	jrp2166@cumc.columbia.edu	Tuesdays, 12 PM – 1 PM	ARB 15th floor conference room (1514)
Alex Perlmutter	asp2183@cumc.columbia.edu	TBD	TBD
Richard Teran	rat2127@cumc.columbia.edu	TBD	TBD

ARB – Allan Rosenfield Building (Mailman School of Public Health)

REQUIRED TEXTBOOK

There is no required textbook for this course. However, there are required article readings each week specified in the course schedule, located in the "assigned readings" folder on Canvas.

RECOMMENDED READING (also see supplemental reading list on last page of syllabus)

For theoretical aspects of epidemiological research and data analytic methods, the following books are also recommended for reading and have been placed on reserve at the Health Sciences library (text also available online through the Health Sciences Library):

1. Rothman K, Greenland S, Lash T. (2008) Modern Epidemiology (3rd edition). Philadelphia: Lippincott-Raven. Referred to as R&G. http://www.columbia.edu/cgi-bin/cul/resolve?clio8363805
2. Hosmer DW, Lemeshow S (2013). Applied Logistic Regression (3nd edition). New York: John Wiley & Sons. http://onlinelibrary.wiley.com/book/10.1002/9781118548387
3. Hosmer DW, Lemeshow S (2008). Applied Survival Analysis. New York: John Wiley & Sons. http://onlinelibrary.wiley.com/book/10.1002/9780470258019

COURSE STRUCTURE

Class will begin with a 1 hour and 50 minute lecture which will be followed by a 50 minute lab session.

COMPUTER LABORATORIES

Laboratories are designed to provide more informal discussions of conceptual issues, provide hands-on experience of statistical modeling and to provide technical assistance to students. Students will be randomly assigned to a laboratory section. All laboratories are held immediately after class from 3:00 PM – 3:50 PM in the following locations:

Laboratory Sections and Locations

- Section 001: Location HSC LL106; T.A. Precious Esie
- Section 002: Location HSC LL107; T.A. Diana Garofalo
- Section 003: Location HSC LL108A/B; T.A. Shadiya Moss
- Section 004: Location HSC LL109A/B; T.A. John Pamplin
- Section 005: Location HSC LL110; T.A. Alex Perlmutter

SAS Installation

Students are expected to bring their laptops to class with SAS installed. *If you do not have SAS installed on your laptop, you will need to let Dr. Genkinger (jg3081@cumc.columbia.edu) and Adiba Ashrafi (ai2337@cumc.columbia.edu) know so that we can make arrangements for you. **You will need SAS to complete laboratories, quizzes and homeworks**.

*Note: the full version of SAS, SAS University, or SAS Studio can be used for this course

SAS Help

Link to SAS Online Tutorials: https://stats.idre.ucla.edu/sas/
Also see the "SAS Help" folder under "Files" in Canvas.

AUDIENCE RESPONSE SOFTWARE

We will be using audience response questions during lecture to encourage participation and to give students an opportunity to participate during lecture. You will receive points for participating and for answering polls correctly. Your responses will count toward 5% of your grade. **You will need to register with Poll Everywhere using your Columbia UNI and password**. Your Poll Everywhere account will be linked to your Courseworks gradebook, therefore, your account **must** be associated with your Columbia email.

If you already have a Poll Everywhere account, verify that it is associated with one of your Columbia email addresses. If you have an account linked to a non-Columbia email (i.e., Gmail, yahoo, Hotmail, etc.), you will need to create a new account. A document will be shared before the first class and posted to Courseworks (Files > Poll Everywhere) detailing the steps necessary to register for a Poll Everywhere account.

To use Poll Everywhere, you can either download the app to your phone or table or access it through the following website: www.pollev.com/jeaninegenki547 . If you would like to use the app on your phone or tablet, please download the app before the first class.

EMAILING

- Email your lab TA and cc: Dr. Genkinger (jg3081@cumc.columbia.edu) and Adiba Ashrafi (ai2337@cumc.columbia.edu) on all emails regarding the course or homework assignments.
- Email Richard Teran (rat2127@cumc.columbia.edu) regarding any Canvas, Poll Everywhere, and lecture recording issues.
- Emails received after 6 PM on the day prior to an assignment's due date may not be answered before the assignment is due.

ASSESSMENT AND GRADING POLICY

Student grades will be based on:

Homeworks..... 20%
Pre-semester assessment.....5%
Quizzes..... 15%
Poll Everywhere participation 5%
Midterm exam 25%
Final exam..... 30%

PRE-SEMESTER ASSESSMENT

Before the start of the semester students must complete the online pre-semester assessment. It is administered through Canvas (Quizzes) and is due before **10:00 AM** on the first day of class. The first part of the assessment reviews epidemiologic concepts from pre-requisite Epi courses and the second part reviews concepts from SAS and Categorical Data Analysis. Review materials are posted on Canvas (Files). The goal of this assessment is to help the instructor and TAs tailor the course and laboratory sessions to students' level of comfort with pre-requisite material. You will be given 1 hour to complete the assessment. The pre-semester assessment is completion based, but students must complete all of it to receive full credit. The assessment is open book; however, students must work ***independently*** to complete the assessment.

HOMEWORK

- Homeworks are due in class at 1:00 PM on Tuesdays before lecture. Please write neatly or type your answers. No credit will be given for illegible answers.
- Late homeworks will ***NOT*** be accepted.
- The lowest homework grade will be dropped so that each homework counts 4% of the total homework grade.
- All students must abide by the Mailman HONOR code.

QUIZZES

Four quizzes (two announced, two unannounced) will be given throughout the semester in either lecture or laboratory sessions. Announced quizzes are take-home and their due dates are listed below in the course schedule. The lowest quiz grade will be dropped so that each quiz counts 5% of the total quiz grade.

MIDTERM EXAM

This exam will be given in-class on **October 15th**. It is a closed-book exam. If there is a valid reason for missing an exam, you must notify the instructor **two weeks prior** to the exam.

FINAL EXAM

The final exam consists of two parts, an **in-class exam on Tuesday, December 10th** (worth 10% of the final exam grade) and a take-home (open-book) portion due Thursday, December 12th (worth 20% of the final exam grade). The take-home portion of the final exam will be distributed via Canvas at the end of day on December 3rd and will be **due in person on Thursday, December 12th at 12:00 PM**. The take-home portion of the exam should be submitted in person and specific submission details will be provided closer to the exam due date. The exam will cover material presented over the entire semester, but it will focus primarily on post-midterm material. The take-home portion of the exam must be turned in on time. For every day late, there will be a 5% reduction in grade.

FINAL GRADE CUTPOINTS:

98.0 or above	A+
94.0 - 97.9	A
90.0 - 93.9	A-
88.0 - 89.9	B+
84.0 - 87.9	B
80.0 - 83.9	B-
78.0 - 79.9	C+

74.0 - 77.9	C
70.0 - 73.9	C-
Below 70.0	F

RE-GRADING POLICY
If students have questions about their grade on any assignment (homework, quiz, exam) they must bring it to Dr. Genkinger.

COURSE WEBSITE
To access the course website please visit https://courseworks2.columbia.edu/courses/88226 and log in with your UNI and UNI password. Our course number is EPIDP8400_001_2019_3. It is essential to visit the course website before each class. It is the primary source of course information, announcements, resources and class assignments, and will be updated regularly. Lecture slides, readings, and exercises will be in PDF format. Please use Internet Explorer, Firefox, or Chrome as your internet browser to access the course website. Computers with Internet access are available at the Health Sciences Library and other locations on campus.

LECTURE RECORDINGS
The class will be taped with the slides. However, this class is not designed as an online course. The taped classes will only be available on a case by case basis if you need to miss a class due to illness etc. or if you have a disability, or other need for audio reinforcement. Please contact Richard Teran (rat2127@cumc.columbia.edu) to request regular access or to request access to the tape for a specific class.

MAILMAN SCHOOL POLICIES AND EXPECTATIONS
Students and faculty have a shared commitment to the School's mission, values and oath.
https://www.mailman.columbia.edu/about/mission-history

Academic Integrity
Students are required to adhere to the Mailman School Honor Code, available online at
https://www.mailman.columbia.edu/people/current-students/community-standards

- Homework assignments:
  - You may discuss concepts and work through homework problems with classmates; however, the work you submit must be your own (e.g., in your own words, performing your own calculations, providing your own code and results).
  - It is in violation of the honor code to use answer keys from previous semesters to complete course assignments. Sharing of answer keys either directly or via internet upload is strictly prohibited in this class.

- Take-home quizzes
  - Quizzes are open book/open notes, but must be completed independently. You cannot discuss the quiz or work with classmates to complete the quiz. You may only discuss the content of the quiz with classmates *after* the quiz due date (*i.e.,* after all students have completed the quiz).

- Final exam:
  - Unlike homeworks, for the final exam you may not discuss concepts or work through exam questions with classmates – you must complete the final exam alone. If you have

any clarification questions after release of the final exam you may direct them to your TA and the professor.

Anyone found to be in violation of these course procedures will be reported to OSA (https://www.mailman.columbia.edu/people/current-students/office-student-affairs).

Disability Access

In order to receive disability-related academic accommodations, students must first be registered with the Office of Disability Services (ODS). Students who have, or think they may have a disability are invited to contact ODS for a confidential discussion at 212.854.2388 (V) 212.854.2378 (TTY), or by email at disability@columbia.edu. If you have already registered with ODS, please make sure you inform Sarah Tooley (st3146@cumc.columbia.edu), the School's liaison to the Office of Disability Services, that you would like to receive accommodations for this class. Then, speak to your instructor to ensure that they have been notified of your recommended accommodations by Sarah Tooley.

Encouraging an Open and Inclusive Classroom Environment

The Department of Epidemiology is committed to creating an educational culture that encourages robust, open, and inclusive classroom environments. Key to achieving this goal is to ensure that all students are included in the conversation and feel comfortable expressing themselves. This commitment is part of our collective enactment of the elements of the Public Health Oath in which we agree to "respect the rights, values, beliefs, and cultures of those individuals and communities with whom [we] work."

An inclusive classroom environment is undermined by microaggressions. Microaggressions are commonplace verbal, behavioral, and environmental indignities, frequently unintentional, that communicate hostile, derogatory, or negative sentiments about individuals on the basis of status characteristics such as race, ethnicity, gender, sexual-orientation, religion, disability, etc. Those who commit a microaggression are usually unaware that they have demeaned another individual, but the consequences for those on the receiving end can be significant. Microaggressions harm individuals by making them feel invalidated, isolated, diminished, and marginalized. They harm the learning environment by making it less inclusive, open and productive.

Recognizing and addressing microaggressions can help mitigate these negative consequences and thereby maintain a robust classroom environment. We believe that responding to microaggressions in the classroom is a critical part of educational growth that leads to a better understanding of the sociocultural issues we seek to investigate as epidemiologists and public health professionals. If you have observed or been the target of a microaggression from a classmate, TA, or faculty member, you are encouraged to bring it to their attention when it happens. Faculty and TAs are willing and prepared to facilitate such engagement, even if they are responsible for the microaggression. If you are uncomfortable speaking up immediately, please contact the TA or instructor outside of class, or another faculty member with whom you feel comfortable, or with any member of the Department senior administration or leadership.

COURSE SCHEDULE

Before class: Pre-semester assessment (Epi concept and SAS review)	
Tues 9/3	**Online assessment (Epi concept and SAS review – available through Canvas)** Learning Objectives: • The first part will contain questions testing basic epidemiology concepts • The second part will contain questions testing basic SAS familiarity • Goal: to help instructors and TAs tailor the course and laboratory sessions to students' level of comfort with pre-requisite material • You will be given 1 hour to complete the quiz which can be taken anytime within the provided 1 week time window
***Pre-semester assessment to be taken on Canvas **by Tuesday, September 3rd at 10:00 AM** (available on Canvas in the Quizzes tab by August 26th)	

Causal Inference, Measures of Effect and Association, Multivariable Model	
9/3	**Lecture 1** Learning Objectives: • Discuss causal inference and the concept of the counterfactual • Distinguish between absolute vs. relative measures of effect and association • Explain the relationships among various measures of association • Discuss the multivariable model and its advantages over tabular analyses for analysis of epidemiologic data
	Lab 1: Let's get reacquainted with SAS: review of OR and RR, and relationships between OR, IRR, and RR
	Reading: Rothman KJ. J Gen Intern Med 2014;29(7):1060-1064.

Precision and Bias	
9/10	**Lecture 2** Learning Objectives: • Discuss and compare precision and bias • Explain how selection bias, information bias, and confounding can influence observed measures of association • Discriminate between selection bias, information bias, and confounding • Distinguish between confounding and mediation
	Lab 2: Bias: assess confounding by stratified analysis
	Reading: Nichol et al. NEJM 2007;357:1373-1381.

Interaction	
9/17	**Lecture 3** Learning Objectives: • Describe and distinguish between statistical and biological interaction • Assess interaction on the additive and multiplicative scales • Explore interaction in stratified and regression analyses
	Lab 3: Interaction: evaluate additive and multiplicative interaction by stratified analysis
	Reading: Flink et al. PLoS ONE. 2013;8(8):e70070
***Homework 1 due	

Case-control Analysis I	
9/24	**Lecture 4** Learning Objectives: • Summarize the design elements of a case-control study • Describe the measure of association used for a case-control study

	• Discuss the fundamental concepts underlying the logistic regression model and when it is useful for the analysis of epidemiologic data • Demonstrate approaches for deciding which covariates to include in a multivariable model
	Lab 4: Logistic Regression: assess confounding and interaction using regression
	Readings: Hogervorst et al. JNCI 2009;101:651-662
*****Homework 2 due**	

Case-control Analysis II	
10/1	**Lecture 5** Learning Objectives: • Demonstrate approaches to evaluate linearity • Evaluate confounding and interaction using a multivariable model • Describe model evaluation • Discuss and apply an extension of the logistic model – Polytomous Regression
	Lab 5: Model Building, Assessing Linearity, Polytomous Modeling
	Readings: Merritt et al. British Journal of Cancer 2014; 110: 1392–1401
*****Homework 3 due**	

Review for Midterm	
10/8	
	Lecture: Midterm review of concepts up through case-control analysis II
	Lab: We will review sample midterm questions in lab based on midterm review problem set
*****Homework 4 due**	

IN CLASS MIDTERM EXAM (CLOSED BOOK)	
10/15	
	In class midterm exam 1:00-3:50pm

Cohort/Follow-up Analysis I & Age, Period, Cohort Effects	
10/22	**Lecture 6** Learning Objectives: • Summarize the design elements of a cohort study • Describe and discuss utility of analysis of age, period, and cohort effects • Describe and demonstrate basic survival analysis: Life Table and Kaplan-Meier methods
	Lab 6: Kaplan Meier / Non-Parametric Survival Analysis
	Readings: Bresalier et al. NEJM 2005;352(11):1092-1102 Keyes et al. IJE 2012;41:495–503

Cohort/Follow-up Analysis II	
10/29	**Lecture 7** Learning Objectives: • Describe and discuss the choices for multivariable models of survival data once we move beyond Kaplan-Meier and Life Table methods • Discuss the Cox Proportional Hazards Model and describe its assumptions • Decide when the Cox model is a good choice for the analysis of epidemiologic data • Demonstrate approaches for evaluating whether the assumptions of this model have been met • Review an application of the Cox model from the literature • Explain the concept of time-varying covariates

	• Demonstrate the utility of time-varying covariates in the analysis of data from cohort studies • Evaluate confounding and interaction using a multivariable model
	Lab7: Cox PH Modeling
	Readings: Tobias et al. NEJM 2014;370:233-44 Powell et al. 2012. (time varying covariates article)

NO CLASS (election day)
11/5
*****Take-home quiz to be taken on Canvas by Tuesday, November 5th at 1:00 PM**

Matched Analysis and Correlated Data	
11/12	**Lecture 8** Learning Objectives: • Describe the purpose and benefits of matching in case-control • Summarize tabular analyses for matched data • Discuss regression models suitable for matched data; apply these methods • Discuss different options for handling correlated data • Compare and contrast GEE and MLMs
	Lab 8: Matching and Modeling: Conditional Logistic Regression
	Readings: Ludvigsson et al. JAMA Psychiatry. 2013;70(11):1224-1230 Shen et al. PLoS ONE 2012; 7(9):e44308
*****Homework 5 due**	

Poisson Regression and Advanced Topics	
11/19	**Lecture 9** Learning Objectives: • Describe and give examples of types of correlated data • Examine commonly used regression models for correlated data • Discuss poisson regression model and describe its assumptions • Decide when the poisson model is a good choice for the analysis of epidemiologic data • Demonstrate approaches for evaluating whether the assumptions of the Poisson model have been met • Describe relative risk regression
	Lab 9: GEE, poisson and relative risk regression
	Reading: Perkins et al. Cancer 2014;120:2514-21 Fang et al. NEJM 2012;366:1310-8.
*****Take-home quiz to be taken on Canvas by Tuesday, November 19th at 1:00 PM**	

Synthesis of the Literature	
11/26	**Lecture 10** Learning Objectives: • Discuss and distinguish between narrative reviews, meta-analysis, and pooled analysis • Explain statistical methods used in meta- and pooled analyses
	Lab 10: Synthesis of Literature and Meta-Analysis
	Readings: Auvinen, Anssi, et al. *Epidemiology* 13.3 (2002): 356-359. Hepworth, Sarah J., et al. *BMJ* 332.7546 (2006): 883-887. Lönn, Stefan, et al. *Epidemiology* 15.6 (2004): 653-659.
	Discussion of Auvinen et al., Hepworth et al., and Lönn et al. during Lab 10
*****Homework 6 due**	

Review for final exam **12/3**
***Take-home portion of final exam available on Canvas by end of day on Tuesday, December 3rd

IN-CLASS FINAL EXAM **12/10**
***FINAL EXAM (IN-CLASS): In-class portion of final exam, December 10th at 1:00-2:50 PM.

TAKE-HOME FINAL EXAM DUE **12/12**
FINAL EXAM (TAKE-HOME): Due on Thursday, December 12th at 12:00 PM. Final exams should be submitted ***in person; submission details will be provided closer to exam date.

(See last page for list of supplemental/recommended course readings)

Epi 3 Supplemental Reading List

(additional supplemental articles may be added throughout the semester)

Lecture	Textbook	Recommended Readings
Lecture 1: Causal Inference, Measures of Effect and Association, Multivariable Model	Rothman (Modern Epi)	Ch 2 (p. 5-13) Ch 3 (p. 32-48) Ch 4 (p. 51-61)
Lecture 2: Precision and Bias	Rothman (Modern Epi)	Ch 9 Ch 10 (p. 148-151, 156-158)
Lecture 3: Interaction	Rothman (Modern Epi) Article (interaction fallacy) Article (interaction tutorial)	Ch 5 Morabia. 1997. J Clin Epidemiol VanderWeele and Knol. 2014 Epidemiol Methods
Lecture 4: Case-control Analysis I	Rothman (Modern Epi) H & L (Logistic) Article (Table 2 Fallacy) Article (confounder control)	Ch 8 Ch 14 (p. 238-253) Ch 15 (p. 274-276) Ch 1 Ch 2 Westreich and Greenland. 2013. AJE McNamee. 2005. Occup Environ Med
Lecture 5: Case-control Analysis II	Rothman (Modern Epi) H & L (Logistic) Article (model assumptions)	Ch 17 (p. 303-305, 321-323) Ch 20 (p. 394-395, 413-415) Ch 3 Ch 8.1 Ch 8.2 Ch 10.8 Barker & Shaw. 2015. AJCN
Lecture 6: Cohort/Follow-up Analysis I & Age, Period, Cohort Effects	Rothman (Modern Epi) H & L (Survival) Szklo & Nieto (A,P,C effects)	Ch 14 (p. 253-257) Ch 15 (p. 273-274) Ch 20 (p. 393-394) Ch 1 Ch 2 Ch 1 (p. 1-14); pdf provided
Lecture 7: Cohort/Follow-up Analysis II	H & L (Survival)	Ch 3

Lecture 8: Matching and correlated data	Rothman (Modern Epi) Vittinghoff (GEE) Website (GEE/repeated measures)	Ch 11 (p. 171-182) Ch 16 (p. 283-288) Ch 21 (p. 434-435) Ch. 7.4-7.6, 7.11; pdf provided Click here
Lecture 9: Poisson and Relative Risk	Rothman (Modern Epi) Vittinghoff (relative risk reg.) Vittinghoff (GLMs/Poisson) Website (MLMs) Website (relative risk regression) Article (relative risk regression) Article (Poisson)	Ch. 14 (p. 240-245) Ch. 5.5.3; pdf provided Ch. 8; pdf provided Click here Click here Zou et al. 2004. AJE Simard et al. 2010. Arch Int Med

H&L(Logistic): Hosmer & Lemeshow "Applied Logistic Regression"
H&L(Survival): Hosmer & Lemeshow "Applied Survival Analysis"
Vittinghoff: Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. Available through the health sciences library (online):
https://clio.columbia.edu/catalog?q=vittinghoff&search_field=all_fields&commit=Search