Graduate Group in Epidemiology and Biostatistics
Introduction to Epidemiologic Research (EPID 7010)
Fall 2024
Syllabus

Overview
This course is intended to provide in-depth, exposure to the theory and methods of epidemiologic research. Topics to be covered include causal inference, measures of disease frequency and association, study design, bias and confounding, validity, and epidemiologic analysis. The course is designed for students entering the PhD program in Epidemiology. However, students from other graduate groups are welcome, as long as they meet the pre-requisites; such students are welcome during any year of study.

Learning outcomes
After completing this course, students will be able to:
- Describe the science of epidemiology
- Demonstrate ability to develop a variety of observational and analytic study designs
- Discuss in detail the quantitative foundations of epidemiologic theory and methods
- Describe current approaches to epidemiologic research
- Identify sources of bias and ways to address them
- Critically assess epidemiologic literature
- Formulate study designs to address specific problems in population health

Class format
This is primarily a lecture course with journal club and laboratory. We will meet every Friday 10:00 to 1:00 in 235 Blockley Hall.

Participating faculty
Course director: John H. Holmes, PhD jhholmes@mail.med.upenn.edu

Course units
This is a 1.0cu course.

Contact hours
The course will consist of one three-hour in-person session each week. In addition, a discussion board will be established on Canvas for out-of-class communication. Office hours are on an ad hoc basis; students should feel free to email me at any time should the need arise.

Course structure
The course is designed around four modules, each focusing on a major area of epidemiologic research principles. Each module (except for the last) consists of a series of sessions which are dedicated to a specific topic or method:
- Basic concepts
- Study design and conduct
- Epidemiologic data analysis
- A special topic to be selected
- Final project presentation, specific to each student's interest area

Class structure
Each class will generally follow the same schedule.
- Hours 1-2: Lecture with participation
- Hours 2-3: Journal club or group discussion: Discussion of key literature pertaining to the week's lecture
- Hour 3: Lab, when scheduled. Lecture adjusted accordingly.

Take-home problem sets associated with each module will be assigned approximately every week and reviewed in the class on the due date.

Course materials
- Required texts (Note: these are freely available online at https://www.library.upenn.edu/vanpelt). Search for the title of each one, select the Online version, and download. These will be in pdf format
  - Lash, VanDerWeele, Haneuse, and Rothman: *Modern Epidemiology*, 4th Edition, 2021.
  - Weiss and Koepsell: *Epidemiologic Methods: Studying the Occurrence of Illness*, 2nd Edition, 2003
  - Carstensen: *Epidemiology with R*, 2021- a copy is also located in *Files | Additional resources*
- Optional:
  - The Epidemiologist R Handbook: https://www.epirhandbook.com/en/index.html
- Course materials on Canvas
  - Materials for each week are provided in the *Files | Class materials* folder, in the subfolder for each class session, labeled as Week *N* (*Date*)
  - Problem set assignments are located in the *Files | Problem Sets* folder, along with a helpful tips and tricks document for attacking the sets
  - The most updated version of this syllabus is located in *Files | Syllabus*

Assignments and evaluation of student performance
The grade for the course will be based on the following:
- **Problem sets (50%).** There are 10 problem sets, each linked to the previous week's material. Each problem set will be worth five points, and the lowest grade will be dropped in tabulating the final grade. Group collaboration on these assignments is highly encouraged, but assignments must be submitted by each student separately written in their own words. In addition to responding to any questions you have about the assignment over email, office hours may be scheduled as needed to discuss any questions about the assignment you may have. *Problem sets are due in Canvas (Assignments folder) no later than 8am the day they are due, as noted in the class schedule below. Problem sets will be discussed in class that day.*

- **Final project (30%).** The final project will propose a study of a problem that you could address using epidemiologic research methods. It will take the following form as a 10 to 12-page report:
  - Statement of the research question
  - Specific aims for the project to address the research question
  - Description of the problem domain, including background and prior work
  - Outline of methods to address the research question
    - Study design
    - Target population
    - Data source(s)
    - Sampling methods (if needed)
    - Analysis plan
  - Consideration of potential limitations of the study
  - References
  - Each student will deliver a 15-minute presentation to the class, with an additional five minutes for questions from the class. Presentations are scheduled for the last day of class, December 13th. PowerPoint slides for presentations or due in the Canvas *Assignments* folder no later than 8am December 13th. **The final report is due on Canvas by 11:59pm, December 20th.**
  - You are expected to provide periodic oral progress reports as the semester proceeds.
  - Group projects
    - Group projects are allowed and encouraged, but the work must reflect the effort of all students in the group. Extra time will be allotted for the presentation, proportional to the number of students in a group.

- **Journal club and class participation (20%).** Journal clubs are an excellent way to explore a topic, exchange ideas, and react to what others (authors as well as your classmates) think about what has been written on that topic.
  - o There are a total of 16 articles we will review this semester. We will review the first two articles together for our first journal club on 9/6 so everyone can see how this is done. After that, we will typically review two articles each week. Each student will be expected to review two articles over the course of the semester. Students will have the option to review both articles in a given week or split their reviews across the semester. A sign-up sheet will be made available during the first two weeks of the semester, so you can select articles which might be a particular interest to you.
  - o Typically, one person will present an article and questions about it to the class. This is an informal presentation, meaning that PowerPoint slides are not required. This presentation should take no more than 15 minutes.
  - o After the presentation, the focus shifts from the presenter to everyone, who will join in the discussion. The discussion will take 15 minutes.
  - o It is assumed that the presenter will have critically appraised the article, but everyone is expected to contribute to the appraisal during the discussion. Thus, it is very important to have read the article before class.
  - o There are two helpful documents in *Files | Journal Club|Journal Club Tips* in Canvas.
  - o The articles to be discussed in a given week's journal club are indicated in the class schedule below, marked with a "$\Rightarrow$" symbol. They are located in Canvas in the *Files|Class materials|Journal Club* folder in the subfolder for each class session, labeled as Week *N* (*Date*). Additional articles for discussion (marked with a black square) are in *Files|Class materials|Readings for discussion.*
  - o The success of this course, and your success in it, is largely dependent on your participation and discussions in class about the course material and anything else that's relevant to the subject. Participation on the discussion board does not contribute to this component of the course grade, however.

A note about academic integrity

First things first
We are focusing our efforts on *mutual discovery* and planting the seeds for future work in epidemiologic research. There will be occasions where we will be dealing with issues and subjects of a sensitive nature. All of us will be learning from each other, with respect for and consideration of our varied backgrounds and lived experiences.

Shared work
Epidemiology is a discipline which is highly collaborative, and there's no better time to start on this than now. You are encouraged to work together on the problem sets and the final project. However, all work that you submit must be your own, and in the case of a final project were you to work in a group, the contribution of each member must be explicitly noted.

Course policy on the use of language model (LLM)-driven chatbots
No large language model (LLM)-driven chatbots, including ChatGPT, will be accepted as a credited author in answer to any work in this course. All cited author attributions included in your work must demonstrate accountability for the work, and LLM tools cannot take such responsibility. As a result, you are not allowed to use, copy (in part or in whole), or cite (in part or in whole) any result from a query posed to a LLM application or website in your answers on problem sets or on your final project. Doing so will be considered plagiarism and a violation of Penn's academic code. That said, you are welcome to use LLM applications to stimulate your thinking about ways to address your work. Please contact me if you have any questions about this policy.

Class schedule

Module	Week	Date	Topic	Readings Journal club: ⇒ Group discussion: ■	Assignment Due
Basic Concepts	1	8/30	• Introduction to course • A brief history • A case study	• Lash:1, 2; Weiss:1, 2	----
	2	9/6	• Measures of occurrence • Critical appraisal • Journal club pointers	• Lash: 4; Weiss: 3, 4, 7 • Armenian • Winkelstein	----
	3	9/13	• Measures of effect and association	• Lash: 5, 18; Weiss: 9 ⇒ Gradus	Problem set 1: Measures of occurrence
	4	9/20	• Causal modeling • Causal inference • Lab 1: Daggity	• Lash: 3; Weiss: 8 ■ Suttorp ■ Goldstein ■ Rand Brothers ⇒ Ioannidis	Problem set 2: Measures of effect and association
Study Design	5	9/27 Online	• Introduction to study designs • Descriptive studies	• Lash: 6, 30; Weiss: 5, 16 ⇒ Haarbauer-Krupa ⇒ Birru	Problem set 3: Causation
	6	10/4	• Cohort studies	• Lash: 7; Weiss: 14 ■ Rieckmann ⇒ Newman ⇒ Nguyen	Problem set 4: Confounding and descriptive study design
	7	10/11	• Case-control studies and variants • Lab 2: Case-control studies	• Lash:8; Weiss:15 ⇒ Kondo ⇒ Linet	Problem set 5: Cohort studies
	8	10/18	• Experimental studies	• Lash:6; Weiss:13 ■ Jackson ■ O'Halloran ■ MRFIT	Problem set 6: Case-control studies
Data Analysis	9	10/25	• Bias, confounding and effect modification • Mediation analysis	• Lash:12, 13, 29 • Weiss:11, 12 ■ VanderWeele ⇒ Hossain ⇒ Mathur	Problem set 7: Experimental studies
	10	11/1	• Interaction analysis • Lab 3: Confounding and effect modification	■ Lash:26, 27 ■ VanderWeele ⇒ Charters ⇒ Cashin	Problem set 8: Bias and confounding
	11	11/8	• Categorical analysis • Risk stratification	■ Lash:17, 19 ⇒ Banack ⇒ Van Rein	Problem set 9: Mediation analysis
	12	11/15	• Regression methods	■ Lash:20, 21; Weiss:12 ■ Nakai	Problem set 10: Stratified analysis

				■ Faust	
		11/22	Thanksgiving holiday		
Data Analysis	**13**	**11/29**	• Longitudinal studies • Time-to-event analysis • Lab 4: Survival analysis	• Lash:22, 24 ⇒ Mitchell ⇒ Lin	----
Special topic	**14**	**12/6**	• Machine learning in epidemiology	■ Bi ■ Du ■ Senaratna	
12/13: Presentations of final projects					