EPID 7020: Advanced Topics in Epidemiologic Research

1. Background and Learning Objectives

The goal of EPID 7020 (*Advanced Topics in Epidemiologic Research*) is to expose students to advanced epidemiologic and statistical research methods and theories that are limitedly or not otherwise covered in curriculum courses; it is the natural successor to EPID 7010 (*Introduction to Epidemiologic Research*, a prerequisite). EPID 7020 is intended for first-year epidemiology PhD students in the Graduate Group in Epidemiology and Biostatistics (GGEB). Non-GGEB graduate group students in any year of study are welcome, if fulfilling the prerequisite and/or at the discretion of the course co-directors. EPID 7020 is built on a framework of the following modules: 1) design; 2) bias; 3) advanced modeling; and 4) translational epidemiology. See §3 below for detail on the sessions constituting each module. EPID 7020 has the following learning objectives: 1) to provide students with an understanding of modern and cutting-edge quantitative methods, advanced topics, and best practices in epidemiologic, statistical, and biomedical research; 2) to develop students' competence and confidence in statistical programing to support accurate and reproducible epidemiologic and biostatistical analyses; and 3) to improve the students' ability to make informed decisions regarding the selection of analytic methods in their individual and collaborative research projects. EPID 7020 emphasizes the following core competencies: 1) knowledge within program area (epidemiologic and biostatistical methods); 2) research skills (study planning, critically appraising published research); and 3) quantitative and computational methodologies (data manipulation, data analysis, statistical coding and debugging, Bayesian inference, data visualization, purposeful statistical inference, and model selection). Through instructor-led lectures, reading of curated tutorials, critical appraisal of papers, and in-class activities (e.g., workshops), EPID 7020 will provide instruction on rigorous and informed epidemiologic study design selection and statistical model selection, estimation, and interpretation.

2. Course Information

2.1 Course Co-Directors

Michael Harhay, PhD, MPH | mharhay@pennmedicine.upenn.edu

Charles Leonard, PharmD, MSCE, FISPE | celeonar@pennmedicine.upenn.edu

2.2 Module Director (in addition to the above-named persons)

Ellen Caniglia, ScD | ellen.caniglia@pennmedicine.upenn.edu

2.3. Session Lecturers

Penn-based (in addition to the above-named persons)
Bryan Blette, PhD | bryan.blette@pennmedicine.upenn.edu
Cheng (Alice) Chen, PhD | cheng.chen@pennmedicine.upenn.edu
Nandita Mitra, PhD | nanditam@pennmedicine.upenn.edu
Alexis Ogdie-Beatty, MD, MSCE | ogdiea@pennmedicine.upenn.edu

External to Penn
(Pending) **Jessie Edwards, PhD** | University of North Carolina | jessedwards@unc.edu
Julia Szymczak, PhD | University of Utah | (new email address forthcoming)

2.4. Location   TBD

2.5. Credits   1.0 course unit

2.6. Prerequisite   EPID 7010 and/or at the discretion of the course co-directors (§2.1). Enrolled students are expected to have prior biostatistics experience or training and knowledge of and/or experience in working in biomedical research or a clinical domain.

2.7. Materials   Textbook(s), biomedical research papers, and assignments, as noted below.

Textbooks   Required | Lash, VanderWeele, Haneuse, and Rothman. Modern Epidemiology, 4th Edition.

Papers   Required | These will generally be uploaded to Canvas by faculty. That said, it is ultimately the student's responsibility to find and read the identified papers. This can often be accomplished via Penn's Biotech Commons biomedical library (https://www.library.upenn.edu/biotech-commons).

2.8. Format   Class will meet each Tuesday from 12:00pm – 3:00pm Eastern. The first session is Tuesday, 01/17/2023.

Lecture   In general, each session will begin with a 60–120-minute faculty-led lecture. Accompanying slide decks, if any, will typically be posted the day prior to the session.

Workshop   In general, each session will end with an activity intended to help students synthesize and/or apply lecture material and/or readings. Activities may range from 15–90 minutes, with the entire session duration never exceeding three (3) hours. Some activities will be faculty-led. Others will be student-led; in fact, some workshops will require a student to lead a comprehensive discussion of a methods topic that extends what is presented by the faculty lecturer. An example of this would be a student led-discussion of the case-time-control design immediately following Dr. Leonard's lecture on self-controlled study designs. At the beginning of the semester, students will be asked to volunteer to fill these slots. As a point of emphasis, regardless of the lead for a given workshop, each student is expected to interact and engage.

Assignments   Approximately six (6) problem sets will be assigned throughout the semester. Unless otherwise instructed, a problem set will be due the second Wednesday after its assignment. For example, if a problem set is assigned on 01/17/2023, it would be due on 01/25/2023. Each problem set will contribute to the final grade (i.e., none are 'dropped'). Collaboration with fellow EPID 7020 students is permissible, but assignments must be separately submitted by each student and written

in her/his own words. Faculty are available to answer questions: a) via email; and/or b) during a to-be-determined office hour.

Please note that there is <u>no</u> reading journal requirement. Furthermore, there is <u>no</u> final project. See the evaluation subsection below for grading details.

Evaluation   Problem sets (summing to 90%); and class participation (quantity and quality), including potentially leading a methods extension presentation (summing to 10%).

3. Session Detail

Module 1: Design | Module director: **Dr. Ellen Caniglia**

Session	Date	Title	Lecturer(s)
1	Tue 1/17	**Longitudinal data collection and clustered study designs: Considerations and interpretations**	**Dr. Harhay**
		Required readings	PMID 7873953, 7703752, 9451271, 15911637, 29905618
		Suggested readings	None
		Workshop activity	Faculty-led
		Problem set	No
2	Tue 1/24	**Self-controlled designs**	**Dr. Leonard**
		Required readings	PMID 24635348
		Suggested readings	PMID 27618829, 24030723
		Workshop activity	**Student-led methods extension** | Case-time control design | At a minimum, should cover PMID 7619931, 9647910 | **TBD**
		Problem set	No
3	Tue 1/31	**Quasi-experimental designs**	**Dr. Mitra**
		Required readings	PMID 33978956, 28239929
		Suggested readings	PMID 27283160, 32879971
		Workshop activity	**Student-led methods extension** | Interrupted time series design | At a minimum, should cover PMID 27283160, 32879971 | **TBD**
		Problem set	Yes | Instrumental variables
4	Tue 2/7	**Target trials**	**Dr. Caniglia**
		Required readings	PMID 26994063, 28748498, 34972229
		Suggested readings	PMID 34942066
		Workshop activity	Faculty-led | What is the target trial? PMID 34972229
		Problem set	Yes | Target trials

Module 2: Bias | Module director: **Dr. Charles Leonard**

5	Tue 2/14	**Sensitivity analysis, secondary analysis, and quantitative bias analysis**	**Dr. Caniglia**

		Required readings	PMID: 25080530, 33778845, 29936049
		Suggested readings	None
		Workshop activity	**Student-led methods extension** | Triangulation, validation studies, <u>or</u> e-values | **TBD**
		Problem set	Yes | **TBD**
6	Tue 2/21	**Measurement error and misclassification**	**Dr. Caniglia**
		Required readings	PMID 15308962, 1591319, 25751609
		Suggested readings	None
		Workshop activity	Faculty-led | Perinatal stressors
		Problem set	Yes | **TBD**
7	Tue 2/28	**Assessing diagnostic tests: Sensitivity, specificity, predictive values, and implications**	**Dr. Ogdie-Beatty & Dr. Chen**
		Required readings	PMID 21548184, 32491423
		Suggested readings	PMID 33057203, 32398230
		Workshop activity	Faculty-led | How good are the diagnostic tests?
		Problem set	No

Module 3: Advanced modeling | Module director: **Dr. Michael Harhay**

	Tue 3/7	Spring term break	
8	Tue 3/14	**Basics of survival analysis and restricted mean survival time**	**Dr. Harhay**
		Required readings	PMID: 29239842, 9836663, 15117797, 9703534, 28546261, 29307954
		Suggested readings	None
		Workshop activity	Faculty-led | Restricted Mean Survival Time estimation
		Problem set	Yes | **TBD**
9	Tue 3/21	**Longitudinal and clustered modeling (technical)**	**Pending**
		Required readings	PMID: **TBD**
		Suggested readings	None
		Workshop activity	Faculty-led | Longitudinal modeling
		Problem set	Yes | **TBD**
10	Tue 3/28	**Mediation**	**Dr. Caniglia**
		Required readings	PMID: 26653405, 27489089
		Suggested readings	None
		Workshop activity	Faculty-led | Mediation analysis for health disparities research
		Problem set	Yes | **TBD**
11	Tue 4/4	**Propensity scores**	**Dr. Leonard**

		Required readings	PMID: TBD
		Suggested readings	TBD
		Workshop activity	**Student-led methods extension** | High-dimensional propensity scores | **TBD**
		Problem set	No
12	Tue 4/11	**Frequentist vs. Bayesian Statistical Analysis**	**Dr. Blette**
		Required readings	PMID: 33270526, 30347031
		Suggested readings	None
		Workshop activity	Faculty-led | Bayesian re-analysis of the ART trial
		Problem set	No

Module 4: Translational epidemiology | Module director: **Dr. Charles Leonard**

		Ontology, epistemology and methodology: Other ways of knowing	**Dr. Szymczak (virtual)**
13	Tue 4/18	Required readings	https://psycnet.apa.org/doi/10.1086/428914 and PMID: 30985531, 27802938
		Suggested readings	None
		Workshop activity	Faculty-led | Researcher reflection
		Problem set	No
14	Tue 4/25	**Knowledge synthesis in the context of missing data, measurement error, and selection bias**	(Pending) **Dr. Edwards (virtual)**
		Required readings	**TBD**
		Suggested readings	**TBD**
		Workshop activity	None
		Problem set	No