

Mel and Enid Zuckerman College of Public Health

University of Arizona

**SYLLABUS**

EPID 677 - Principles of Genetic Association Studies

# Spring 2022

**Time**: Mondays, 9:00 – 11:50 am

**Modality**: Drachman, A122

**Instructor(s) and Contact Information**:

Yann Klimentidis, PhD

Associate Professor of Public Health

Department of Epidemiology and Biostatistics

[yann@email.arizona.edu](mailto:yann@email.arizona.edu)

**Instructor Availability**:by appointment – please email me.

**Course Description**: The course will focus on the principles, methods, and challenges involved in dissecting the genetics of human complex traits and diseases using association studies. It will consist of a theoretical component, a hands-on applied component (using R software for genetic data analysis), as well as discussion, group presentations, and group activities. Specific topics will include: heritability, population genetics and population stratification, epidemiological design strategies for genetic association studies, GWAS, genetic correlation, gene-by-environment interactions, polygenic risk scores and genetic risk prediction, Mendelian randomization, and genomic and other omic databases.

**Course Prerequisites**: EPI573A and EPI576A, or some background in epidemiology and/or statistics are strongly suggested, but not required.

**Course Objectives**:

To provide an overview of the concepts, methods, and hands-on applications to design, conduct, and interpret genetic association studies and to conduct genotype analyses. During this course, students will analyze genotype and phenotype data, read and critically evaluate genetic epidemiology literature, and present in both written and oral formats on genetic epidemiology studies.

**Expected Learning Outcomes:**

**MPH Competencies Covered:**

1. Search, describe and summarize findings from the scientific literature to describe the epidemiology of a public health problem, identify health disparities and identify risk factors.

2. Compare the relative strengths and weaknesses of epidemiological study designs, and choose the most appropriate design for specific research questions.

3. Calculate and interpret appropriate measures of disease frequency and excess risk across multiple study designs.

4. Assess and identify strategies to minimize bias in analytic, along with assessing effect modification and confounding, then stratifying or adjusting as appropriate in analyses.

**MS Competencies Covered:**

1. Critique and synthesize appropriate literature and research findings to address a research question.

2. Identify potential sources of bias for various study designs and their impact on study quality.

3. Conduct descriptive and analytic analyses, including strategies to assess confounding and effect modification methods, to make statistical inferences.

4. Demonstrate ability to manage and analyze epidemiological data from a variety of sources.

5. Organize and deliver clear presentations of research findings in varying professional formats to diverse audiences.

**PhD Competencies Covered:**

1. Evaluate the integrity, comparability, and limitations of data to make inferences related to analyses and results.

2. Lead group interactions competently, ethically, respectfully and professionally to diverse audiences.

3. Organize and deliver clear presentations of research findings in varying professional formats to diverse audiences.

**Course Notes**:

Lecture presentations and reading material will be posted on D2L.

**Required Texts or Readings**:

There are no required textbooks. Required readings will be posted on D2L.

**Required Software:**

For this course, you will need to download R Studio and R software, and be able to run these softwares on your computer. R Studio is basically a ‘nice place’ in which to code and run R software. Here are links for each of these:

<http://archive.linux.duke.edu/cran/>

<https://rstudio.com/products/rstudio/>

**Course Requirements**:

Class attendance; read assigned paper and complete paper dissection questions prior to class; co-lead paper presentation/discussion; take quizzes on lecture material; completion of in-class R exercises

**Quizzes:** We will have a quiz on most lecture/discussion days (i.e. not data analysis week). This quiz will cover the material from the previous lecture. You will take the quiz on D2L on your laptop or phone. I will allow you to drop your lowest grade.

**Article dissections:** These will be the same for each paper that we discuss on lecture weeks. They will be due BEFORE class. You will be asked to identify the hypotheses and predictions, describe the methods, results, and conclusions, and provide your assessment of the paper/study. I expect about a 2-3 sentence response per question. Your grade will be based on answering the question correctly, original and creative thought on your part, level of critical evaluation and evidence of reading the paper thoroughly. I will allow you to drop your lowest grade.

**Group presentation & leading discussion:** The first week of class you will sign up for a week to lead a paper discussion. Your role for these discussions is not to talk the whole time, but to provide some background and then pose thoughtful questions that can be discussed in class. Each member of the group should present/lead discussion an equal amount.

**Groups should email the instructor indicating the paper that they have chosen on the Monday of the week prior to your presentation.**

For background: do not summarize the article or go into much depth about what was in the article (everyone should have read it). Instead try to put the article into a greater context. What do we already know about the general topic? Why is it interesting? Briefly describe what other studies have found (a good place to find references to other studies is in the introduction where the background is discussed-look up these papers and read the abstract, look at the figures, maybe skim through the methods).

Discussion: try to start with the basics. For example, ask the class what they thought about the rationale and methods. Ask them if they agree with the conclusions and what they thought about the results, etc. Then try to have 3-4 additional questions designed to promote discussion in class. Debates are a good idea. You should also copy and paste figures from the results section onto slides for discussion.

Power point: make a short power point presentation to help you lead discussion. It should include:

-2-5 slides on background

-a slide for each table and figure (just copy and paste these in)

-a slide for each of your discussion questions

Presentation suggestions:

• First provide a brief summary of background information that is necessary to understand the paper.

• Potential questions to pose when leading discussions:

• Why did the authors conduct this study?

* How does this study contribute to our knowledge of genetic epidemiology?

• What is the rationale for conducting this study?

• Why did the authors choose this particular population?

• What are the main findings? (here it might be a good idea to go present figures).

• How do these findings relate to the hypotheses/research topic?

• What aspects of the study did the authors deal with well? Why do you think that was a good way to handle the research topic?

• Are any parts of the study controversial? Explain.

• How could the study be improved?

• What are some future studies that could be done along these lines/on this topic?

**Your grade will be based on how closely you follow the above guidelines and suggestions, how much critical and original content you present, and the quality of the slides in terms of how effective they are at communicating your points. It will also be based on how equally you distribute effort in presenting among the 2-3 presenters.**

**R exercises**: You are expected to complete R-based exercises in which you will analyze genetic and phenotypic data. These in-class R exercises must be turned in at the end of class, unless otherwise notified. The lowest in-class exercise will be dropped.

**Grading:**

in-class R exercises: 25%; Discussion lead: 15%; Quizzes: 25%; Participation/Group activities: 15%; Paper dissections: 20%.

Grading scale: 100-90: A; 80-89: B; 70-79: C; 60-69: D; below 60: F. This scale may be revised by instructor, if necessary.

**Communications**: You are responsible for reading emails sent to your UA account from your instructor and the announcements that are placed on the course web site. Information about readings, news events, your grades, assignments and other course related topics will be communicated to you with these electronic methods. The official policy can be found at: <https://www.registrar.arizona.edu/personal-information/official-student-email-policy-use-email-official-correspondence-students>

**UA Smoking and Tobacco Policy:**The purpose of this Policy is to establish the University of Arizona’s (University) commitment to protect the health of University faculty, staff, students, and visitors on campuses and in its vehicles. The official policy can be found at: <http://policy.arizona.edu/ethics-and-conduct/smoking-and-tobacco-policy>

**University Course Policies:** (please see the following URL):<https://academicaffairs.arizona.edu/syllabus-policies>

**Plagiarism:**  What counts as plagiarism?

* Copying and pasting information from a web site or another source, and then revising it so that it sounds like your original idea.
* Doing an assignment/essay/take home test with a friend and then handing in separate assignments that contain the same ideas, language, phrases, etc.
* Quoting a passage without quotation marks or citations, so that it looks like your own.
* Paraphrasing a passage without citing it, so that it looks like your own.
* Hiring another person to do your work for you, or purchasing a paper through any of the on- or off-line sources.

**Syllabus Changes:** Information contained in the course syllabus, other than the grade and absence policies, may be subject to change with reasonable advance notice, as deemed appropriate by the instructor.

**COURSE SCHEDULE** (may be revised):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Date | Activity | Topic | Assignment due at the end of class | Quiz | Discussion paper |
| 1 | January 24 | Data analysis | Syllabus, introduction, introduction to R and genetic data |  |  |  |
| **2** | January 31 | Lecture & Discussion | Basic molecular genetics / Heritability |  |  | TBD |
| **3** | February 7 | Data analysis | Basic genetic dataset analysis in R | Exercise 1 |  |  |
| **4** | February 14 | Lecture & Discussion | Population genetics / Linkage disequilibrium |  | Quiz #1 | TBD |
| **5** | February 21 | Data analysis | Online genetic databases; data quality control; population stratification | Exercise 2 |  |  |
| **6** | February 28 | Lecture & Discussion | Genetic association studies I |  | Quiz #2 | TBD |
| **7** | March 14 | Data analysis | Statistical tests for genetic association | Exercise 3 |  |  |
| **8** | March 21 | Lecture & Discussion | Genetic association studies II |  | Quiz #3 | TBD |
| **9** | March 28 | Data analysis | Statistical tests for genetic association, cont. | Exercise 4 |  |  |
| **10** | April 4 | Lecture & Discussion | Gene-by-Environment interactions & PRS |  | Quiz #4 | TBD |
| **11** | April 11 | Data analysis | Genetic risk scores (aka, polygenic risk scores) | Exercise 5 |  |  |
| **12** | April 18 | Lecture & Discussion | Mendelian randomization |  | Quiz #5 | TBD |
| **13** | April 25 | Data analysis | Mendelian randomization | Exercise 6 |  |  |
| **14** | May 2 | Lecture & Discussion | Misc.: prediction, sequencing, bio- annotation, LD-score regression,PheWas |  | Quiz #6 | TBD |