NUTR 319: Intermediate Epidemiology

Spring 2024

TIME: Thursday @ 9am-12pm (note time change for 1/25/2024 and 4/18/2024)

LOCATION: Jaharis 105. Please bring your laptop to class each week and have SAS installed on your laptop. You can also access SAS via TTS virtual lab (https://access.tufts.edu/tts-virtual-lab-vdi) but we ask you to test virtual lab access prior to the first class.

FACULTY:
Fang Fang Zhang, M.D., Ph.D.
The Neely Family Professor and Associate Professor
Friedman School of Nutrition Science and Policy
Office: Jaharis 250
Phone: 617-636-3704
Email: fang_fang.zhang@tufts.edu
Office Hour: By appointment

PREREQUISITES: NUTR 204, NUTR 206, and NUTR 309 or equivalents, or concurrently taking NUTR 309 or equivalents

COURSE CREDIT: 1 Credit Hour (3 SHUs)

COURSE GOALS:
Intermediate Epidemiology exposes students to a variety of key concepts and methods when carrying out epidemiologic studies and teaches students applied skills in analyzing epidemiologic data and interpreting study findings appropriately.

This course includes a 2-hour lecture session followed by a 1-hour lab session. The lecture session will present epidemiologic methods and concepts beyond the Principles of Epidemiology and review relevant statistical methods and their applications in epidemiologic studies. The lab session will prepare students with practical skills in conducting and analyzing epidemiologic studies using SAS. The lab session will be taught in a computer lab equipped with SAS.

After completing the course, students should be expected to demonstrate mastery in the following areas:

1. Study design
2. Measures of disease occurrence
3. Measures of association
4. Precision and validity
5. Confounding
6. Mediation
7. Effect measure modification
8. Bias
9. Causal inference
10. Communicating scientific results

Students should also be expected to acquire the following data analysis skills:

1. To perform descriptive analysis, produce $2 \times 2$ table, and calculate odds ratio (OR) and relative risk (RR);
2. To evaluate dose-response relationship, select between continuous and categorical variables, and determine cut-points for categorical variables;  
3. To estimate sample size and statistical power, and perform sensitivity analysis in power calculation;  
4. To assess Kappa as a reliability measure, and evaluate the consequence of misclassification;  
5. To determine whether a variable fulfills the criteria for a confounder, and control confounding through stratification and adjustment;  
6. To examine effect modification, and understand the difference in the evaluation of confounding, mediation and interaction;  
7. Perform linear regression and logistic regression, and evaluate confounding and effect modification in regression models;  
8. Perform conditional logistic regression in matched studies and polytomous logistic regression for polytomous outcomes;  
10. Perform Cox proportional hazard and Poisson regression in cohort studies

This is an option to fulfill a required course for the Epidemiology and Biostatistics concentration in the Tufts MPH program. This course delivers concentration competencies required for the MPH degree. To pass this course, you must successfully complete each competency-based assignment. If you do not get a passing grade on the competency-based assignment, please arrange to meet with me to discuss next steps so we can ensure you attain the competency.

<table>
<thead>
<tr>
<th>Competency</th>
<th>Sessions Where Competency is Taught</th>
<th>Competency Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop a causal framework using epidemiologic tools to conceptualize and operationalize the occurrence and determinants of a public health problem.</td>
<td>Lectures 1, 2, 4, 5, 6, 8, 9, and 10</td>
<td>Weekly Homework</td>
</tr>
<tr>
<td>Critique epidemiologic evidence and available data in the context of a public health problem.</td>
<td>All Lectures</td>
<td>Mid-term and Final Exam</td>
</tr>
</tbody>
</table>

**TEXTBOOKS:**

Recommended:

  - Links to eBook: [https://tufts-primo.hosted.exlibrisgroup.com/permalink/f/14dinuo/01TUN_ALMA51231669940003851](https://tufts-primo.hosted.exlibrisgroup.com/permalink/f/14dinuo/01TUN_ALMA51231669940003851)

  - Links to eBook [https://tufts-primo.hosted.exlibrisgroup.com/permalink/f/14dinuo/01TUN_ALMA21106086940003851](https://tufts-primo.hosted.exlibrisgroup.com/permalink/f/14dinuo/01TUN_ALMA21106086940003851)


Additional readings will be assigned. These readings will be available on reserve in the library and, if available in electronic format, on the course web site.
COURSE MATERIAL:
Lecture notes, lab exercises, homework and data sets used for lab and homework exercises will be posted on the class website at Canvas. Students are expected to print and bring lecture notes to each lecture, print and bring lab exercises and download and bring data sets to each lab session.

CLASSROOM FORMAT:
The primary format of this course consists of weekly sessions that include a 2-hour lecture given by the instructor and a 1-hour lab. The total length of one session is 3 hours although the exact length of each lecture and lab could vary week by week. All lectures will be recorded; recording will be available on Canvas shortly after the class.

Class discussion is strongly encouraged. Please try not to feel intimidated by the size of the class or the nature of the material. The only bad question is the one not asked. If something puzzles you, the chance is excellent that the same thing puzzles several of your classmates. We cannot answer any questions you do not ask.

HOMEWORK:
Each week students will be asked to perform certain steps of analysis on topics that are covered in the lecture and lab sessions using the datasets provided. These homework exercises will be graded and you will receive credit for completed exercise (3 points for each completed homework; 1.5 points for each partially completed homework; and 0 point for late homework). Students can consult Dr. Zhang and/or the Teaching Assistant regarding problems you face with these analyses. All homework assignments must be typed and submitted online at the Canvas website by Thursday 9 AM (before the class starts).

COURSE GRADE:
Course grade is based on the following:

- 30% Homework (3 pts each completed homework, 1.5 pts each partially completed homework, 0 pts for late homework or not handed in)
- 10% Lab and class participation
- 20% Mid-term exam
- 40% Take-home final

ALL ARE WELCOME
We believe that the diversity of student experiences and perspectives is essential to the deepening of knowledge in this course. We consider it part of our responsibility as instructors to address the learning needs of all of the students in this course. We will present materials that are respectful of diversity in race, color, ethnicity, gender, age, disability, religious beliefs, political preference, sexual orientation, gender identity, socioeconomic status, citizenship, language, or national origin among other personal characteristics. We would like to create a welcoming classroom environment in which all feel comfortable to participate and learn.

ACCOMMODATION OF DISABILITIES
Tufts University is committed to providing equal access and support to all students through the provision of reasonable accommodations so that each student may access their curricula and achieve their personal and academic potential. If you have a disability that requires reasonable accommodations please contact the Friedman School Assistant Dean of Student Affairs at 617-636-6719 to make arrangements for determination of appropriate accommodations. Please be aware that accommodations cannot be enacted retroactively, making timeliness a critical aspect for their provision.
Week-by-week Course Learning Objectives

Week 1
Lecture 1: Causality in Epidemiology and Study Design Strategies
Learning Objectives of Lecture 1:
◼ To review five important concepts in epidemiologic studies including study design, confounding, bias, effect modification and causal inference

Lab 1: Data Cleaning, Editing and Summarization using SAS
Learning Objectives of Lab 1:
☐ To review basic SAS procedures for data cleaning and editing
☐ To perform descriptive analysis

Homework Distributed: Homework 1: Descriptive analysis
Homework Due: None
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics. Chapter 1, and Chapter 10 Section 10.2, pp.376-392

Week 2
Lecture 2: Precision and Validity
Learning Objectives of Lecture 2:
◼ To understand which factors determine power and sample size
◼ To understand how selection bias occurs and how it affects measures of association
◼ To understand the consequence of differential and non-differential misclassification

Lab 2: Sample size, power, and misclassification
Learning Objectives of Lab 2:
☐ To estimate power and sample size for epidemiologic studies using SAS
☐ To evaluate the consequence of misclassification

Homework Distributed: Homework 2: Sample size and power, measurement errors
Homework Due: Homework 1
Reading Assignment:
(1) Szklo & Nieto Epidemiology beyond the Basics. Chapter 4

Week 3
Lecture 3: Measures of Disease Occurrence and Association
Learning Objectives of Lecture 3:
◼ To understand the relation among different measures of disease occurrence (cumulative incidence, incidence rate, incidence density, prevalence and odds)
◼ To understand the relation among different measures of association (risk ratio, attributable risk, odds ratio and rate ratio and risk/rate difference)

Lab 3: Estimating Measures of Association from Two-by-two Tables
Learning Objectives of Lab 3:
☐ To produce measures of association (risk ratio and odds ratio) using two-by-two tables
☐ To investigate further the relationship among different measures of association
☐ To learn when and how to perform ANOVA and chi-square analysis using SAS
Homework Distributed: None
Homework Due: Homework 2
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics. Chapters 2 and 3

Week 4
Lecture 4: Confounding
Learning Objectives of Lecture 4:
- To understand criteria to evaluate confounding
- To learn ways to control for confounding using stratification
- To understand the difference between confounding and mediation

Lab 4: Evaluation of confounding in stratified analysis, difference between confounding and mediation
Learning Objectives of Lab 4:
- To determine whether a potential confounder fulfills the criteria for being a confounder
- To evaluate confounding in a stratified analysis and produce weighted summary (M-H)
- To understand the difference between confounding and mediation

Homework Distributed: Homework 3: Confounding
Homework Due: None
Reading Assignment:
(1) Szklo & Nieto Epidemiology beyond the Basics. Chapter 5

Week 5
Lecture 5: Heterogeneity of Effects: Effect Modification/Interaction
Learning Objectives of Lecture 5:
- To learn how to evaluate interaction in stratified analysis on additive and multiplicative scales
- To understand the difference between interaction on an additive scale and interaction on a multiplicative scale
- To understand interaction fallacy
- To understand the difference among confounding, mediation and effect modification

Lab 5: Evaluation of effect modification/interaction in stratified analysis
Learning Objectives of Lab 5:
- To evaluate additive and multiplicative interaction in stratified analysis using SAS
- To explore interaction fallacy
- To understand the difference in the evaluation of confounding, mediation and effect modification

Homework Distributed: Homework 4: Effect modification/Interaction
Homework Due: Homework 3
Reading Assignment:
(1) Szklo & Nieto Epidemiology beyond the Basics. Chapter 6
* Midterm review questions distributed

Week 6
Lecture 6: More on Confounding, Mediation and Effect Modification
Learning Objectives of Lecture 6:
- To know the difference in the causes and consequences of confounding, mediation and effect modification and how to handle them appropriately in statistical analysis
No Lab.
Homework Due: Homework 4
*Midterm study questions distributed

Week 7
Mid-term

Week 8
Lecture 7: Stratification and Regression Modeling
Learning Objectives of Lecture 7:
- To understand the difference between stratification and regression modeling
- To learn different forms of regression modeling in epidemiologic studies, their assumptions and when to use them

Lab 6: Linear Regression
Learning Objectives of Lab 6:
- To learn basic SAS procedures for simple and multiple linear regression
- To evaluate confounding and effect modification/interaction in linear regression
- To evaluate linearity assumptions and modeling of non-linear relations with linear models

Homework Distributed: Homework 5: Linear regression and linearity assumption
Homework Due: None
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics Chapter 7, pp.227-258 and pp.274-279

Week 9
Lecture 8: Case-control Studies and Data Analysis in Case-control Studies I
Learning Objectives of Lecture 8:
- To understand how different types of control sampling relate to the measures of association estimated in case-control studies
- To understand the difference between the nested case-control design and the case-cohort design and the advantages and disadvantages of these designs
- To understand the basic procedures for logistic regression modeling

Lab 7: Logistic Regression in Case-control Studies
Learning Objectives of Lab 7:
- To perform logistic regression analyzing data collected from case-control studies
- To assess confounding and interaction in logistic regression models

Homework Distributed: Homework 6: Logistic regression modeling in case-control studies
Homework Due: Homework 5
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics Chapter 7, Section 7.4.3., pp.258-265
Optional reading: Hosmer & Lemeshow 2000 Applied Logistic Regression (Chapter 1) – library reserve

Week 10
Lecture 9: Case-control Studies and Data Analysis in Case-control Studies II
Learning Objectives:
- To learn the principles of choosing between a categorical or continuous variable for exposure in case-control studies
- To learn when to apply and how to use polytomous/multinomial logistic regression for polytomous outcomes in case-control studies

Lab 8: Dose-response relationship and polytomous logistic regression
Learning Objectives of Lab 8:
To evaluate dose-response relationship in logistic regression
To perform polytomous logistic regression
To understand the difference among linear regression, logistic regression and polytomous logistic regression

Homework Distributed: Homework 7: Dose-response relationship and polytomous logistic regression
Homework Due: Homework 6
Optional reading: Rothman, Greenland & Lash, Modern Epidemiology, Section on Expansion of Logistic Models, pp.413-415– library reserve

Week 11
Lecture 10: Matching
Learning Objectives of Lecture 10:
- To understand the advantages and disadvantages of matching
- To understand the differences between matching in a cohort study and matching in a case-control study

Lab 9: Conditional logistic regression in matched case-control studies
Learning Objectives of Lab 9:
- To learn to conduct matched data analysis

Homework Distributed: Homework 8: Conditional logistic regression in matched case-control studies
Homework Due: Homework 7
Reading Assignment: Rothman, Greenland & Lash, Modern Epidemiology Chapter 11

Week 12
Lecture 11: Cohort studies and Data Analysis in Cohort Studies I
Learning Objectives of Lecture 11:
- To learn steps to conduct cohort studies and understand the difference between open cohort and closed cohort studies
- To understand how to estimate measures of associations in cohort studies
- To understand the fundamentals of survival analyses

Lab 10: Life-table and Kaplan-Meier Methods
Learning Objectives of Lab 10
- To learn basic SAS procedures for life-table and Kaplan-Meier methods in cohort studies

Homework Distributed: Homework 9: Life-table and Kaplan-Meier methods
Homework Due: Homework 8
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics. Chapter 2, pp. 50-57

Week 13
Lecture 12: Cohort Studies and Data Analysis in Cohort Studies II
Learning Objectives of Lecture 12:
- To understand the difference between non-parametric, semi-parametric and parametric approaches to survival analysis in cohort studies
- To understand the fundamentals of Cox Proportional Hazard Modeling
- To learn how to apply and when to use Poisson regression in cohort studies

Lab 11: Cox Proportional Hazard Regression and Poisson Regression
Learning Objectives of Lab 11:
- To perform Cox Proportional Hazard Model in cohort studies
- To examine the assumption of proportional hazard
- To perform Poisson regression in cohort studies
Homework Distributed: Homework 10: Cox proportional hazard model and Poisson regression
Homework Due: Homework 9
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics. Chapter 7, Sections 7.4.4 and 7.4.5, pp. 265-270
Optional reading: Hosmer & Lemeshow 2000 Applied Survival Analysis (Chapter 1-4) – library reserve
* Final review questions distributed and take-home final distributed

Week 14
Lecture 13: Communicating Epidemiologic Findings and Final Review
Reading Assignment: Szklo & Nieto Epidemiology beyond the Basics. Chapter 9
Homework Due: Homework 10

Week 15
Take-home Final
## Intermediate Epidemiology Course Outline

<table>
<thead>
<tr>
<th>Week #</th>
<th>Date</th>
<th>Lecture Topic</th>
<th>Lab Topic</th>
<th>Homework Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>1/18/2024</td>
<td>Lecture 1: Causality in Epidemiology and Study Design Strategies</td>
<td>Lab 1: Data Cleaning, Editing and Summarization</td>
<td>None</td>
</tr>
<tr>
<td>Week 2</td>
<td>1/26/2024 (Fri) @1-4pm *note time change</td>
<td>Lecture 2: Precision and Validity</td>
<td>Lab 2: Sample size, Power and Misclassification</td>
<td>Homework 1</td>
</tr>
<tr>
<td>Week 3</td>
<td>2/1/2024</td>
<td>Lecture 3: Measures of Disease Occurrence and Association</td>
<td>Lab 3: Estimating Measures of Association from Two-by-two Table</td>
<td>Homework 2</td>
</tr>
<tr>
<td>Week 4</td>
<td>2/8/2024</td>
<td>Lecture 4: Confounding</td>
<td>Lab 4: Evaluation of Confounding in Stratified Analysis</td>
<td>None</td>
</tr>
</tbody>
</table>

### 2/22/2024 NO CLASS – MONDAY CLASS SCHEDULE on THURSDAY

<table>
<thead>
<tr>
<th>Week 6</th>
<th>2/29/2024</th>
<th>Lecture 6: More on Confounding, Mediation and Effect Modification</th>
<th>No Lab</th>
<th>Homework 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 7</td>
<td>3/7/2024</td>
<td><strong>In-Class Mid-Term (3 Hours)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 8</td>
<td>3/14/2024</td>
<td>Lecture 7: Stratification and Regression Modeling</td>
<td>Lab 6: Linear Regression</td>
<td>None</td>
</tr>
</tbody>
</table>

### 3/21/2024 NO CLASS - SPRING BREAK

<table>
<thead>
<tr>
<th>Week 9</th>
<th>3/28/2024</th>
<th>Lecture 8: Case-control Studies and Data Analysis in Case-control Studies I</th>
<th>Lab 7: Logistic Regression in Case-control Studies</th>
<th>Homework 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 10</td>
<td>4/4/2024</td>
<td>Lecture 9: Case-control Studies and Data Analysis in Case-control Studies II</td>
<td>Lab 8: Dose-response Relationship and Polytomous Logistic Regression</td>
<td>Homework 6</td>
</tr>
<tr>
<td>Week 11</td>
<td>4/11/2024</td>
<td>Lecture 10: Matching</td>
<td>Lab 9: Matched Data Analysis (Student Presentations #1)</td>
<td>Homework 7</td>
</tr>
<tr>
<td>Week 12</td>
<td>4/19/2024 (Fri) @1-4pm *note time change</td>
<td>Lecture 11: Cohort Studies and Data Analysis in Cohort Studies I</td>
<td>Lab 10: Lifetable and Kaplan-Meier Methods</td>
<td>Homework 8</td>
</tr>
<tr>
<td>Week 13</td>
<td>4/25/2024</td>
<td>Lecture 12: Cohort Studies and Data Analysis in Cohort Studies II</td>
<td>Lab 11: Cox Proportional Hazard Regression and Poisson Regression (final exam and final review questions distributed)</td>
<td>Homework 9</td>
</tr>
<tr>
<td>Week 14</td>
<td>5/2/2024</td>
<td>Lecture 13: Communicating Epidemiologic Findings and Final Review</td>
<td>No Lab (Student Presentations #2)</td>
<td>Homework 10</td>
</tr>
<tr>
<td>Week 15</td>
<td>5/9/2024</td>
<td><strong>Take-Home Final Exam Due</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>