Overview and Objectives: This course is designed for medical researchers who are not biostatistics majors. The course will focus on the basic concepts and the use of logistic regression models. At the end of this course, a student should be able to: 1) recognize the multivariable problem addressed by logistic regression in terms of the types of variables considered; 2) assess the relationship of risk factors and a binary outcome using a logistic regression model; 3) compute and interpret estimated risk of developing a disease or an odds ratio from a fitted logistic regression model; 4) understand the limitation in the estimation when different study designs are used; 5) develop analytic skills through the analysis of data sets taken from the fields of medicine and public health; and 6) develop oral and written communication skills through the description of analytic strategies and the summarization and interpretation of results.

Responsibilities:

- There will be reading assignments in the textbook. All readings are expected to be completed before class.
- Students will be assigned three homework assignments that will be graded. All homework assignments will be assigned a due date. You are encouraged to discuss and work together on problems, but you must write up your results individually, i.e. very similar papers will not be accepted. Homework assignments are to be turned in at the beginning of class on the due date. No assignments will be accepted via email. Late homework assignments will be penalized 10% per day past the due date (unless prior arrangements have been made with Dr. Park) until the following lecture, at which no late homework’s will be accepted.
- There will be a course quiz during the second-to-last class consisting of short answer and multiple-choice questions. The focus of this quiz will be on understanding 1) what to do and when to do it, 2) how to interpret what you did, and 3) more focused questions on the important issues. I will give more specifics as the class progresses.
- There will be a final project distributed on the last day of class and will be due exactly one week later. This project must be turned in on time to receive full credit, and will be subject to the same late penalties as the homework (outlined above). Students will be asked to analyze a data set and answer specific questions pertaining to this analysis and its results. Students must work on the final project independently. Discussion, consulting, or working with other people is not permitted.

Course Requirements:

- Homework assignments 45%
- Final project 30%
- Course quiz 25%
**Attendance Policy:**

- Students are expected to sign-in to each class (computer provided in suite lobby). If a problem is encountered with the sign-in system, please contact the course instructor(s) as well as Allie Giel (giela@upmc.edu) immediately.

**Course Grading Scale:**

For the computation of the final course grade as well as for the course assignments, the following grading scale will be used:

- 90 - 100 = A  
- 80 - 85 = B  
- 70 - 75 = C  
- 60 – 65 = D  
- 86 – 89 = B+  
- 76 – 79 = C+  
- 66 – 69 = D+  
- < 60 = F

**Required Textbook(s):**


**Supplemental Textbook(s):**


**Website resources:** All homework assignments, course information, and communication will be available at [http://courseweb.pitt.edu](http://courseweb.pitt.edu).

**Academic Integrity:**

Students in this course will be expected to comply with the University of Pittsburgh’s Policy on Academic Integrity ([http://www.provost.pitt.edu/info/ai1.html](http://www.provost.pitt.edu/info/ai1.html)). Any student suspected of violating this obligation for any reason during the semester will be required to participate in the procedural process, initiated at the instructor level, as outlined in the University Guidelines on Academic Integrity. This may include, but is not limited to, the confiscation of the examination of any individual suspected of violating University Policy. Furthermore, no student may bring any unauthorized materials to an exam, including dictionaries and programmable calculators.
Course Schedule

Session 1: Introduction to Logistic Regression

At the conclusion of this lecture, the student will be able to:
1. Explain when logistic regression could be used in research

Topics:
1. Course overview
2. Introduction to logistic regression
3. Hypothesis testing
4. Interpretation of the logistic regression model
5. Prediction of a binary outcome

Required Reading(s):
1. Vittinghof et al., Section 5.1 (Single Predictor Models)

Homework assignment 1:
1. See Courseweb for handout.

Competencies
1. Problem Formulation: Critically review published studies that use various research methodologies and identify possible sources of bias and potential health disparities therein.
2. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

Session 2: The Multiple Logistic Regression Model

At the conclusion of this lecture, the student will be able to:
1. Describe measures of association for logistic regression models
2. Describe how multiple logistic regression can be used to address questions of confounding and interactions

Topics:
1. Introduction to multiple predictor variables
2. Interpretation of odds ratio
3. Interpretation of predictions and risk ratios
4. Fitting and interpreting interaction variables
5. Issues in coding covariates

Required Reading(s):
1. Vittinghof et al., Section 5.2 (Multipredictor Models) – 5.2.4 (Interaction)

Competencies
1. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.
Session 3: Likelihood Ratio Test & Model Selection

At the conclusion of this lecture, the student will be able to:
1. Contrast model selection techniques for different research questions

Topics:
1. Likelihood ratio test: definition, interpretation, implementation in STATA
2. Strategies for model selection

Required Reading(s):
1. Vittinghof et al., Section 10.2 - 10.4

Homework assignment 2:
1. See Courseweb for handout.

Due Today:
1. Homework assignment 1

Competencies
1. Problem Formulation: Critically review published studies that use various research methodologies and identify possible sources of bias and potential health disparities therein.
2. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

Session 4: Model Diagnostics and Propensity Score Analysis

At the conclusion of this lecture, the student will be able to:
1. Determine logistic regression model validity
2. Explain when propensity score analysis could be used in research
3. Describe different propensity score adjustment methods

Topics:
1. Checking basic model assumptions
2. Evaluating overall model fit
3. Evaluating the fit of individual observations: outliers, leverage, influence
4. Examining multicollinearity
5. Introduction to propensity score analysis
6. Estimating propensity score
7. Using estimated propensity score to reduce bias

Required Reading(s):
1. Vittinghof et al., Section 5.4 (Checking Model Assumptions and Fit)

Competencies
1. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.
Session 5: Prediction and Classification

At the conclusion of this lecture, the student will be able to:
1. Describe statistical techniques for assessing diagnostic accuracy

Topics:
1. Basic classification measures
2. Receiver operator characteristic curves

Required Reading(s):
1. Vittinghof et al., Section 5.2.5 (Prediction) – 5.2.6 (Prediction Accuracy)

Homework assignment 3:
1. See Courseweb for handout.

Due Today:
1. Homework assignment 2

Competencies
1. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

Session 6: Case-control Studies / Multinomial & Ordinal Regression

At the conclusion of this lecture, the student will be able to:
1. Compare and contrast logistic regression models for case-control studies in the matched and unmatched settings

Topics:
1. Overview of case-control and cohort study designs
2. Interpretation of the odds ratio and relative risk
3. Interpretation of probabilities
4. Matched case-control studies
5. Conditional logistic regression
6. Multinomial and ordinal logistic regression models

Required Reading(s):
1. Vittinghof et al., Section 5.3 (Case-control Studies) & 5.5.6 (More Than Two Outcome Levels)

Competencies
1. Problem Formulation: Critically review published studies that use various research methodologies and identify possible sources of bias and potential health disparities therein.
2. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.
3. Data Management and Biomedical Informatics: Organize datasets (variable display and structure) appropriately for given statistical techniques.
4. Methodology: Design basic features of research protocols based on specific research questions, appropriately addressing bias.
5. Written Communication: Organize and report statistical results.
Session 7: Correlated Binary Data / Course Quiz

At the conclusion of this lecture, the student will be able to:
1. Describe statistical techniques for study designs with correlated observations

Topics:
1. Correlation between binary variables
2. Within subject correlation structures
3. Generalized estimating equations (GEE)
4. Generalized linear mixed models (GLMM)
5. Course quiz

Required Reading(s):
1. Vittinghof et al., Section 7

Due Today:
1. Homework assignment 3

Competencies
1. Problem Formulation: Critically review published studies that use various research methodologies and identify possible sources of bias and potential health disparities therein.
2. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.
3. Written Communication: Organize and report statistical results.

Session 8: Course Quiz Review / Final Project Details

At the conclusion of this session, the student will be able to:
1. Demonstrate a full understanding of the appropriate use and limitations of logistic regression models.

Topics:
1. Review Answers for Course Quiz
2. Overview of Final Project Details

Final Project:
1. See Courseweb for handout.

Competencies
1. Applied Analytical Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.