LEARNING OBJECTIVES
BIOST 540
LONGITUDINAL AND MULTILEVEL DATA ANALYSIS
SPRING 2018

Introduction:

In public health and medicine correlated data arise in two common scenarios: the natural grouping of observations in "clusters"; and the collection of observations that are taken as repeated measures on individual units over time. Longitudinal data analysis refers to the methods that are appropriate for repeated measures data. Multi-level analysis methods refer to techniques that characterize systematic and random variation when data have a hierarchical (nested) organization. Although longitudinal and multi-level data involve correlated observations, the scientific goals associated with each data structure are quite different. Therefore, this course will first focus on longitudinal data where the element of time is crucial for the formulation of mean models and for characterization of random variation (correlation). We will then focus on issues and methods for data that come in clusters, and data that may involve multiple levels of clustering.

Background:

Upon entering this course you are expected to have completed courses in introductory statistics or biostatistics, multiple regression, elementary categorical and censored survival data analysis, and multiple logistic regression analysis of data from epidemiologic case-control studies. You should know how to fit multiple linear and logistic regression models, and how to perform hypothesis tests about regression coefficients. You should be familiar with case-control, cross-sectional, and cohort study designs.

Objectives:

After completing this course the student can ordinarily expect to:
1. Describe the impact of correlation on statistical estimates of uncertainty (standard errors).
2. Formulate an analysis plan for data that are collected at baseline and at a single follow-up time.

3. Formulate an analysis plan for repeated measures data when the goal is to compare different groups over time.
4. Describe the types of correlation models that are appropriate for both longitudinal data and for clustered, or multi-level data.
5. Describe the impact of missing data, and to formulate analysis plans that attempt to correct for such selection bias. In addition, students should recognize the key assumptions required for the analysis of missing data, and appreciate the role for sensitivity analysis.