Chapter 1

Introduction to Longitudinal Data Analysis

Longitudinal studies are designed to measure intra-individual change over time. Repeated observations are made on individual subjects, usually at a set of common time points specified by the study protocol. A main objective of longitudinal studies is to relate change over time in individuals to their characteristics (exposure, sex, etc.), or to an experimental condition (drug treatment arm, time since baseline, etc.). In some studies, exposures or experimental conditions may change during the course of the study (as in crossover designs or repeated measures experiments). For example, in the typical crossover design each subject receives every treatment in sequence, with suitable washout periods prior to each treatment. The sequence of treatments is determined a priori by randomization. In repeated measures studies, each subject is measured under a set of pre-specified conditions; differences in the response due to conditions are of primary interest. Outcomes may be measurements, counts, or dichotomous indicators, and we may have multivariate outcomes measured at each of several occasions as well.

In the ideal setting, we will have all subjects measured at the same set of occasions; this greatly facilitates the analysis and interpretation. Some studies may be unbalanced by design, as, for example, when measurements are costly and/or invasive, so only a subset of subjects are measured at all occasions. In other instances it may be very difficult to obtain measurements on all subjects on the same set of occasions. This is especially true when studying human subjects over a long period of follow-up, and when studying clinic populations where illness is a big factor in patient availability. Observations may be mistimed and/or missing,