Clinical research investigators must all address the four classic elements of design before they can move forward with confidence: Study Type, Study Design, Study Time-frame, and Source for Data.

- **Study Type** is the way to distinguish between a descriptive study and a hypothesis-driven study. Classical scientific method teaches about hypothesis-driven research, but it is important to acknowledge the important role of descriptive studies in clinical research. Case studies and case series are descriptive as well as many of the most important papers on natural history of disease (e.g., Change in CD4 cell counts among a cohort of HIV positive, ART initiators). For hypothesis-driven studies the most common issues to address are the logic, detail, and direction of the hypothesis. Vague or general study questions are not enough for the testing of a specific hypothesis, so the time, energy, and help one puts into the development of the hypothesis can yield a stronger study for the investigator.

- **Study Design** refers to the standard designs defined in epidemiology, specifically Case-series, Cross-sectional, Case-control, Cohort, and Experiment. The definitions of these designs are based on understanding the sampling of subjects and collection of information with complete regard to the timing of exposures and disease as well as the inclusion criteria for the study. A strong understanding of these designs will allow the investigator to identify the strongest design for the specific question and resources available.

- **Study Time-frame** is a reference to calendar time (as opposed to disease time stated above). If all the data to be used in the study already exist (e.g., clinic database), then the time frame is retrospective. If the data to be used in a study have yet to be observed/recorded, then the time frame is prospective. Note that the time frame is a separate issue from study design. That is, a cohort study can be a retrospective study or a prospective study, depending on when the observations were made relative to ‘today’.

- **Source for Data** has a strong role in understanding the impact of your findings. Using existing data such as existing electronic medical records or a publicly available dataset (e.g., NHANES) requires the investigator to change his or her hypothesis to match the information gatherer. And, most commonly, the investigator will have to deal with a large amount of data that is missing or in error, all leading to less-than-rigorous conclusions. The most ideal source is the enrollment of subjects from whom observations are collected via well-defined and perfectly followed study protocol. This reduces the number of errors and amount of missing data and allows for the protocol to maintain uniform methods of obtainment and measurement throughout the collection of data.

Patrick Tarwater, PhD
Foster SOM El Paso Campus-Chief of the Division of Biostatistics and Epidemiology