BASIC TERMINOLOGY

- Variables

- Categorical
- Continuous

- Outcome / Endpoint
- Predictor

- Covariates

- Independent
- Dependent

OBJECTIVE

To review common statistical terminology associated with definition of a research question

PURPOSE

In order to design a clinical trial (or any research or experiment), you need to define the question and then define the variables needed to answer that question statistically

COMMON VARIABLES

- **Continuous:** age (in years), weight, height, BMI, systolic blood pressure, disease specific scales like the DAS28, Timed 25 Foot Walk
- **Categorical:** Treatment Groups, Likert Scales, Event Status, Sex, Race, Age (in groups)

VARIABLES

VARIABLE: A variable is any characteristics, number, or quantity that can be measured or counted. A variable may also be called a data item

Variable Types:

- **Continuous:** a numeric variable
  - In theory, have an infinite number of possible values
  - Has a unit of measure, for example: age in years, pounds
  - May have a lower limit (floor) or upper limit (ceiling)
- **Categorical:** mutually exclusive groups
  - Has a limited number of levels or options
  - Could be intervals from a continuous variable, such as age in groups: <20, 20-24, 25-29, 30+
  - Distance between levels may not be measurable or consistent spaced, for example: none, some, many or blonde, brunette, red, grey
  - May be ordered (ordinal): low, medium high
  - May not have an inherent order (nominal): Black, White, Multi-race, Other race
  - Dichotomous: a categorical variable with only 2 levels
    - Yes/No, Event/No Event, Exposed/Not Exposed

VARIABLES IN STATISTICAL USE

OUTCOMES: variables that are measured for change or occurrence

- Expected to change or be altered due to a treatment, intervention or exposure
- Also called the Endpoint, Response, Dependent, or Y variable
- Appears in the left-hand side of a modeling equation, y =

PREDICTORS: variables that are generally fixed (do not change)

- Variables used to predict the outcome
- In a clinical trial, often called the main predictor and is usually the intervention (treatment) or exposure
- Also called the Independent or X variable
- Appears in the right-hand side of a modelling equation, y = x

COVARIATES: a type of predictors that is NOT of primary interest but including it would make the model more accurate

- Often measured at the start of the study before the intervention, i.e. baseline measures
- Should be selected before, *a priori*, the trial begins
- Used in Adjusted models
- Appears in the right-hand side of a modelling equation, y = x1 + x2 + x3… where x1 is the main predictor of interest