LOGISTIC REGRESSION

- Continuous Outcomes

- Main Effect
- Covariates

- Probability
- Likelihood

- Nominal
- Ordinal

- Log Odds
- Odds Ratio
- Logit

OBJECTIVE

To identify key concepts associated with logistic regression

PURPOSE

To review appropriate use & interpretation of logistic regression

COMMON VARIABLES

- **Categorical outcomes:** Event/No Event, Success/Failure, Mild/Moderate/Severe, Tumor Type (I, II, III), etc.
- **Main Effect:** In clinical trials, the main predictor (effect) is the focus of the study, e.g. treatment arms, intervention arms, etc.
- **Any type of covariate variables:** continuous or categorical variables for adjustment, e.g. sex, race, age, disease stage, etc.

LOGISTIC REGRESSION

Logistic regression is the most commonly used method when modelling categorical outcomes as a function of treatment or intervention, with or without adjusting for relevant covariates.

- **Simple Logistic Regression:** categorical outcome with predictor.
- **Multiple (Covariate Adjusted) Logistic Regression:** Treatment group as the main effect, with adjustment for covariates.
- **Nominal vs Ordinal Logistic Regression:** Nominal has only 2 groups in your outcome (Yes or No) or there is no inherent order on your outcome (Agree, Neutral, Disagree). Ordinal has a rank order to the responses (Low, Moderate, High).
- **Probability:** The number with the event in the treatment arm out of the total number in that treatment arm.
- **Odds vs Odds Ratio:**
  - The odds give the likelihood of the event happening compared to it not happening

$$odds\ A = \frac{probability\ of\ the\ event\ for\ Group\ A}{1 - (probability\ of\ the\ event\ Group\ A)}$$

  - The odds ratio (OR) gives the odds of the event happening for Group A compared to Group B

$$odds\ ratio = \frac{odds\ A}{odds\ B}$$

- **Covariate Adjusted Logistic Regression:** When specifying a set of covariates, the OR can be given for treatments. ORs can also be given for covariates, e.g. are the odds higher for Males vs Females.

INTERPRETATION & CONSIDERATIONS

Goal: The goal with Logistic regression is to determine:

- How likely the outcome of interest is to occur for each treatment (%)
- The odds of having the outcome of interest in one treatment compared to another (Odds Ratio)

Interpretation: The primary focus is to answer "Is the outcome of interest more/less likely in one Treatment Group compared to another?"

- Graphs here should be either the % of each response or a forest plot (OR with 95% CI) NOT the regression plots (they don't make sense to non-statisticians)
- Report the Odds Ratios with 95% CI and a p-value, if applicable

Assumptions: Logistic regression relies on the underlying assumptions of linear regression by using what is called the Logit transformation as the outcome, i.e. the mathematical log of the odds of an event happening

$$logit(p) = \log\left(\frac{p}{1-p}\right) = log(odds)$$

Why is this important? If treating the outcome in this manner is not appropriate, then neither are the logistic regression results.

Bottom line: Use a statistician when you need logistic regression.