Sample Statistical Analysis Plan for Maternal Smoking and Infant Health

Please note that this is not the only Statistical Analysis Plan possible for this example. This is merely an example (others may be equally valid).

Primary Objective:

- Estimate the effect of maternal smoking on birth weight for healthy single-birth infant males.

Secondary Objectives:

- Estimate the effects of gestation, mothers' age, parity, mothers' height and mothers' pre-pregnancy weight on birth weight for healthy single-birth infant males.

Data Collection:

- Data from the Child Health and Development Studies (CHDS between 1960 and 1967 among women in the Kaiser Foundation Health Plan in Oakland, California. The Kaiser Health Plan is a prepaid medical care program. The women in the study were all those enrolled in the Kaiser Plan who obtained prenatal care in the San Francisco-East Bay area and who delivered at any of the Kaiser hospitals in Northern California.
- Female births, children with birth defects/serious genetic illnesses, children who died before 28 days after birth and twin/multiple births removed from consideration

Variables Under Consideration:

- **Primary objective & secondary objective outcome:** Birth weight (in pounds)
- **Primary/secondary objective covariates**: mother's smoking status, gestation (duration of pregnancy in days since first day of the last normal menstrual period), parity, mother's age (in years), height (in inches) and pre-pregnancy weight (in pounds).

Missing Data Procedures:

- Cases without birth weight or smoking status information are to be removed from the analysis
- Cases with less than 50% of other covariates missing to have missing data imputation performed

Summaries to be presented:

- Boxplots of birth weight for smoking and non-smoking mother groups on same plot
- Boxplots of birth weight for mothers with no previous births and mothers with previous births

- Scatterplot of birth weight (y-axis) versus gestation, mothers' age, parity, mothers' height and mothers' pre-pregnancy weight, with different symbols for smoking/non-smoking mother cases
- Descriptive statistics by smoking status of birth weight.

Models to be fitted:

- *Primary objective:*
  Linear Regression with the following mean model
  *Birth Weight = intercept + ß1*smoking status*

  If residuals suggest non-homoscedasticity or skewness, outcome variable changed to log(birth weight) and same model fit again

  or

  Two sample T-test comparing
  H0: Mean BW for non-smokers = Mean BW for smokers
  Vs
  H1: Mean BW for non-smokers is not = Mean BW for smokers

  Perform test at a 5% significance level and assess normality assumptions.

- *Secondary Objective:*
- *Linear Regression with the following mean model*
  *Birth Weight = intercept + ß1*smoking status + ß2*gestational age + ß3*parity indicator + ß4*mother's age + ß5*mother's height + ß6*mother's pre-pregnancy weight + possible interactions*

  If residuals suggest non-homoscedasticity or skewness, outcome variable to changed to log(birth weight) and same model fit again. Working at a 5% significance level perform variable selection to assess whether interactions or covariates can be removed from the model.