Hypothesis: (null) There is no difference in the ability of aeromedical crewmembers to perform intubations in the BK-117 and EC-135 helicopters.

10 crewmembers enrolled in the study. Randomized each crewmember into two groups: Group A and Group B for the BK-117. Each crewmember performs the skill three (3) times in each setting for an (n) 30 per group.

Group A
  Intubate in BK-117, then intubate in controlled setting

Group B
  Intubate in controlled setting, then intubate in BK-117

Re-randomize crewmembers and utilize the EC-135 aircraft. Each crewmember performs the skill three (3) times in each setting for an (n) of 30 per group.

Group C
  Intubate in EC-135, then intubate in controlled setting

Group D
  Intubate in controlled setting, then in EC-135

Thus a total of 120 intubations in the BK-117, EC-135, and controlled setting.

Is this a valid process in evaluating the hypothesis? Are there enough numbers (n)? What can we do to make this a better study? What methods can we use to evaluate this data? ANOVA, Chi-square test to determine statistical significance, t-test, etc.