Statistics

Sample size

$$n \sim \frac{2(z1\text{-}\alpha/2 + z1\text{-}\beta)^2 \sigma^2}{\Delta^2}$$

- n is the sample size
- α is the significance level - often set at 0.05 so we accept a 5% chance of making a type I error
- 1-β is the power – often set at 0.8 so we accept an 80% chance of avoiding a type II error
- Δ is the effect size
- σ is the variance within the population

Type I error – states there is a difference when no difference is found (usually set at 0.05 – therefore 5% chance of type I error)

Type II error – states no difference when there is a difference (power (1-β) set at 80-90% therefore 10-20% chance of type II error)

		Disease			
		+ve	-ve		
Test	+ve	a	b	a + b	PPV a/a+b
	-ve	c	d	c + d	NPV d/c+d
		a + c	b + d	a + b + c + d	
		Sensitivity a/a+c SNOUT	Specificity d/b+d SPIN		Accuracy a + d /a+b+c+d

+ve LR = sens/1-spec High is better: >2 good, >20 great
-ve LR = 1-sens/spec Lower is better: <0.5 good, <0.05 great

ROC (receiver operator curves)
Graphical display of how good a diagnostic test is
Relationship between sens & 1-spec at different threshold values
Area under curve - %age of times you get it correct

ANOVA – analysis of variance (regression methods & looking at areas under curve)

Correlation – conceptually used when interested in the closeness of the relationship between two variables

Regression – establish cause and effect