STATISTICS AND DEFINITIONS

Mean – mathematical average
Median – midpoint of range
Mode – most frequent value in the range

Above are the same in normal distribution.

Variance – describes the dispersion of the measures in a given population.

Standard Deviation – the square root of the variance, another measure of dispersion.

Standard error of the mean – inversely proportional to the square root of the sample size, i.e. gets smaller as the sample gets bigger. Used to compute CI's.

Parametric Tests – data is from a normal distribution, e.g. t test and Pearson's r. More powerful than non-parametric.

Non-parametric Tests – data is from any distribution, e.g. X2, Mann-Whitney U-test, Spearmman's rank co-efficient.

	GOLD STD POS	GOLD STD NEG
INDEX TEST POSITIVE	**A** (TRUE POSITIVE)	**B** (FALSE POSITIVE)
INDEX TEST NEGATIVE	**C** (FALSE NEGATIVE)	**D** (TRUE NEGATIVE)

Sensitivity - % of people **with** the disease who are test positive.
True positive rate. How good your test is at finding the disease. **A/A+C.**
SNOUT – A negative result of a sensitive test rules the disease **out.**

Negative predictive value - % with a negative test who do not have the disease. Post-test probability of a negative test. D/C+D. Rule out test. Good if prevalence of disease is low.

Specificity - % of people **without** the disease who are test negative, or how well the test detects the absence of the disease.
True negative rate. How good your test is at excluding the disease. **D/D+B.**
Spin – a positive result of a specific test rules the disease **in.**

Positive predictive value - % with a positive test who have the disease. Chance that a positive test result % will be right. Post-test probability of a positive test. **A/A +B.** Rule **in** test. Good if prevalence of a disease is high.

100% sensitive = 100% NPV.

Pre-test probability. Prevalence of disease in given population.

Likelihood ratio – How the result of a test changes the pre-test probability, or how much more or less likely the diagnosis after testing.
"Likelihood that a given test result would be expected in a patient with the target disease compared to the likelihood of the same result in a patient without."

Post-test odds – Pre-test odds x likelihood ratio.

Positive likelihood ratio – How much more likely you are to have a positive test result if you have the disease than if you do not, (sensitivity/1-specificity)(>2.0 good).

Negative likelihood ratio – How much more likely you are to have a negative test result if you do not have the disease than if you do. (1-sensitivity/specificity). Neg LR zero = sensitivity 100% (<0.5 good).

THEREFORE!!
Low risk and negative sensitive test = very low risk.

Type 1 error – To say there is a difference when there isn't (Null hypothesis true, but reject it). **α (alpha) –** the type 1 error rate, usually 5%, p=0.05.

Type 2 error – To say there isn't a difference when there is. (fail to reject null hypothesis). **β (beta) –** the type 2 error rate, usually 10 – 20%.

Power – 1-β. Takes into account the CI you are willing to accept.

Confidence interval – A measure of the precision of your result compared to the population as a whole. 95% C1 = 1:20 chance that population (or true) value lies outside of your result.

Odds ratio – A way of expressing the good that treatment, or harm that exposure, might cause.

$$\frac{\text{odds in treated or exposed group}}{\text{odds in control group}}$$

OR > 1 = harm for exposure
2 = good for treatment

Internal validity – acceptable level of bias in the study.

External validity – generalization of the result.

Receiver operator characteristic (ROC) curve – graphical display of how good a diagnostic test is. x axis is 1-specificity, y axis is sensitivity.
Area under the curve is the % of time you get it correct.

P value (0.05) Probability.

If there is truly no difference between these results, we would only find this result (or wore) on <5y.

(?Odds ratio) % disease in treated e.g. NTD 3.3%
Relative Risk % disease in control 7.8% = 0.42

Diagnostic odds ratio = true/false = (a * d)/(b * c)

Relative risk reduction – Proportion reduction in disease between the treated and the untreated groups.

$$\frac{\text{expected disease} - \text{observed disease}}{\text{expected disease}}$$

Absolute risk reduction - % expected - % observed (or) (expected – observed?)
100

Number needed to treat – Number of patients that intervention needs to be given to so 1 life is saved or 1 outcome achieved.

$$\frac{1}{\text{absolute risk reduction \%}}$$

e.g. 20% mortality in control group, 15% mortality in treated group.
RRR = 20-15 = 0.25, (25%)
20

ARR = 20-15 = 0.05, (5%)
100

1 NNT = 1 = 20
ARR 0.05

Or 20% + control, 0% + Px
RRR = 20.0 = 1 Therefore only if 100%
20 & ARR 1.0, is
NNT = 1
ARR = 20.0 = 2 or 20%
100

NNT = 5

Kappa – measure of agreement beyond chance
Good > 0.7

“chance – corrected” agreement.

Valid measurements → do they measure what they are supposed to
Reliable measurements → are they reproducible.

Sensitivity Analysis → In meta analysis, results of all studies, then only PRCT’s etc.