statistics

biostatistics

Biostatistics

Type I Error (alpha): A difference between study and control groups is found when in reality there is none. Standard = 5%.

Type II Error (beta): No difference between study and control groups is found when in reality there is a difference. Standard = 20%.

types of data

Types of Data

Nominal: Numbers are arbitrary.

Ordinal: Numbers denote rank order only.

Interval: Numbers denote units of equal magnitude and rank order.

Parametric: Interval data in a normal distribution.

Standard Deviation (SD): Measure of the scatter of data in a normally distributed sample; 95.44%of the data will fall within 2 SD of the mean.

SD = square root of the variance.

Standard Error of the Mean (SE): SE = SD/√n. Used to calculate confidence intervals but not a measure of scatter. Should not be used in place of SD.

Confidence Interval (CI): The estimated range of values likely to include the true value for the entire population. The standard is 95%.

Power Calculation (1 - β): Statistical power is the ability of an experiment to find a significant difference between groups when in fact one exists.

Note: As a increases, so does power. As n is increased, β decreases and power increases; that is, the chance of either a type I or type II error is reduced.

Intention-to-Treat Analysis: All data are analyzed according to what group the subject was assigned to regardless of what treatment the subject actually received: analyzed as randomized

study design

Study Design: The research methodology used. There are basically four categories. From weakest to strongest, these are:

1. Case series.
2. Case-control study.
3. Cohort study.
4. Randomized clinical trial.

two by two table

Two-by-Two Table:

		Disease/Outcome	
		+	-
Test or Exposure	+	a	b
	-	c	d
	Total	a + c	b + d

diagnostic tests

For diagnostic tests

Sensitivity: Probability that the test will be (+) when the disease is present. a/a + c

Specificity: Probability that the test will be (-) when the disease is absent. d/b + d

Positive Predictive Value: Probability that the disease is present given a (+) test. a/a + b

Negative Predictive Value: Probability that the disease is absent given a (-) test. d/c + d

exposure or therapy

For association (with exposure or therapy)

Relative Risk (RR): Estimates the magnitude of an association between exposure and disease (or in the case of therapy, the negative association between treatment and morbid outcome). The relative risk indicates the likelihood of development of disease in the exposed group relative to those who were not exposed (also called risk ratio).

$$RR = \frac{\text{Incidence in exposed group}}{\text{Incidence in unexposed group}} = \frac{a/(a+b)}{c/(c+d)}$$

Relative Risk Reduction (RRR): Expressed as a percentage reduction in events in treated versus untreated groups.

$$RRR = \{1 - [a/(a+b)]/[c/(c+d)]\} \times 100\%$$

Odds Ratio (OR): For case-control studies, RR cannot be used because participants are selected on the basis of disease, not exposure. The RR can be estimated by the OR, however.

$$OR = \frac{a/c}{b/d} = \frac{ad}{be}$$

Attributable Risk (AR): A measure of association that provides information about the absolute effect of the exposure or the excess risk of disease in those exposed compared with those unexposed.

$$AR = (\text{Incidence in exposed group}) - (\text{Incidence in unexposed group}) = [a/(a+b)] - [c/(c+d)]$$

Absolute Risk Reduction (ARR): A measure of the treatment effect. Note the order is reversed compared with AR.

$$ARR = [c/(c+d)] - [a/(a+b)]$$

Number Needed to Treat (NNT): The inverse of the ARR: 1/ARR.

$$NNT = 1/[c/(c+d)] - [a/(a+b)]$$