LITERATURE REVIEW/STATISTICAL SUMMARY

VA Connecticut Healthcare System
Office of Public Health, VACO
Yale University School of Medicine

CLCW SME training August, 2017

Confidence Intervals (CI)

- A confidence interval calculated for a measure of treatment effect (exposure) shows the range with the true treatment (Exposure) effect is likely to lie (subject to a number of assumptions)

- Confidence intervals are preferable to p-values, as they tell us the **range of possible effect sizes** compatible with the data

- Confidence intervals that has the value of no difference (crossing 0) indicates that the treatment (exposure) under investigation is not significantly different from the control.

What is p-value ($p < 0.05$)

- A *p-value* is calculated to assess whether trial results are likely to have occurred simply through chance (assuming that there is no real difference between new treatment and old, and assuming, of course, that the study was well conducted)

- P-values simply provide a cut-off beyond which we assert that the findings are "statistically significant *(i.e. p<0.05)*

Box 1. Hypothesis testing and the generation of p-values

The logic of hypothesis testing and p-values is convoluted. Suppose a new treatment appears to outperform the standard therapy in a research study. We are interested in assessing whether this apparent effect is likely to be real or could just be a chance finding: p-values help us to do this.

In calculating the p-value, we first assume that there really is no true difference between the two treatments (this is called the **null hypothesis**). We then calculate how likely we are to see the difference that we have observed just by chance if our supposition is true (that is, if there is really no true difference). This is the p-value.

So the p-value is the probability that we would observe effects as big as those seen in the study if there was really no difference between the treatments. If p is small, the findings are unlikely to have arisen by chance and we reject the idea that there is no difference between the two treatments (we reject the null hypothesis). If p is large, the observed difference is plausibly a chance finding and we do not reject the idea that there is no difference between the treatments. Note that we do not reject the idea, but we do not accept it either: we are simply unable to say one way or another until other factors have been considered.

But what do we mean by a 'small' p-value (one small enough to cause us to reject the idea that there was really no difference)? By convention, p-values of less than 0.05 are considered 'small'. That is, if p is less than 0.05 there is a less than one in 20 chance that a difference as big as that seen in the study could have arisen by chance if there was really no true difference. With p-values this small (or smaller) we say that the results from the trial are statistically significant (unlikely to have arisen by chance). Smaller p-values (say $p<0.01$) are sometimes called 'highly significant' because they indicate that the observed difference would happen less than once in a hundred times if there was really no true difference.

Important points to remember!

- Statistical significance does not necessarily mean that the effect is real: by chance alone about **one in 20 significant findings will be spurious.**
- Statistically significant does not necessarily mean clinically important. It is the size of the effect that determines the importance, not the presence of statistical significance.
- Non-significance does not meant “no effect”. Small studies will often report non-significance even when there are important, real effects which a large study would have detected.

Summary of Effect Measures

Table 1. Summary of effect measures

Measure of effect	Abbreviation	Description	No effect	Total success
Absolute risk reduction	ARR	Absolute change in risk: the risk of an event in the control group minus the risk of an event in the treated group; usually expressed as a percentage	ARR=0%	ARR=initial risk
Relative risk reduction	RRR	Proportion of the risk removed by treatment: the absolute risk reduction divided by the initial risk in the control group; usually expressed as a percentage	RRR=0%	RRR=100%
Relative risk	RR	The risk of an event in the treated group divided by the risk of an event in the control group; usually expressed as a decimal proportion, sometimes as a percentage	RR=1 or RR=100%	RR=0
Odds ratio	OR	Odds of an event in the treated group divided by the odds of an event in the control group; usually expressed as a decimal proportion	OR=1	OR=0
Number needed to treat	NNT	Number of patients who need to be treated to prevent one event; this is the reciprocal of the absolute risk reduction (when expressed as a decimal fraction); it is usually rounded to a whole number	NNT=∞	NNT=1/initial risk

What is Bias?

- **Bias** is term that covers any systematic errors resulting form the way the study was designed, executed or interpreted.
  - Lack of or failure in) randomization, leading to unbalanced groups
  - Poor blinding, leading to unfair treatment and assessments
  - Large no. of patients lost to follow-up