Worksheet: Finding Associations in the Data

Relative Risk

This is the ratio of the rate (or probability) of an event in an "exposed" group, to the rate of the same event in an "unexposed" group, typically used in cohort studies.

For example: Do depressed residents make more medication errors? One hundred psychiatry residents were followed for one month to determine whether a diagnosis of depression ("exposure") led to more medication errors ("event rate").

Of the 100 residents, 25 met criteria for depression and 75 did not. Six of the depressed residents made an error (event rate = 6/25 = 0.24), while 10 of the non-depressed residents made an error (event rate = 10/75 = 0.13).

$$RR = \frac{\text{event rate in people exposed}}{\text{event rate in people NOT exposed}} = \frac{6/25}{10/75} = \frac{0.24}{0.13} = 1.8$$

Thus, depressed residents had an 80% (or 1.8 times) higher risk of making an error than non-depressed residents (Fahrenkopf AM et al, *BMJ* 2008;336:488).

Odds Ratio

This is the ratio of the odds that an outcome will occur given a particular exposure, to the odds of the outcome in the absence of the exposure. Most commonly used in case-control studies.

For example: Is borderline personality disorder (BPD) predictive of recurrent suicidality in adolescents evaluated for suicidal behavior? In a case-control study, 77 adolescents (cases) with recurrent suicidal behavior within a six month period were compared with 186 adolescents (controls) without recurrent suicidality. A total of 205 had a diagnosis of BPD ("exposure"), 70 of whom presented with recurrent suicidality ("outcome") while 135 did not. Of the remaining 58, only 7 had recurrent suicidality.

$$OR = \frac{\text{odds of outcome in people exposed}}{\text{odds of outcome in people NOT exposed}} = \frac{70/135}{7/51} = \frac{0.518}{0.137} = 3.8$$

Thus, the odds of recurrent suicidal behavior in the six months after an initial onset of suicidal behavior are 3.8 times greater for adolescents with a diagnosis of BPD than those without (Greenfield B et al, *J Can Assoc Child Adol Psychiatry* 2008;17(4):197–201).

If the "event" is rare, then the odds ratio is a good approximation of relative risk, but if the event is common, the odds ratio will overstate the risk (see Sedgwick P, *BMJ* 2014;348:g1407)

Statistical Significance of RR and OR

Relative risks and odds ratios are usually reported with a "confidence interval" (CI), which shows the range of ratios in which the *actual* ratio is likely to lie, with some high degree of certainty (usually 95%). In the Greenfield et al study mentioned previously, for instance, the OR was reported as: 3.8, 95% CI: 1.6–8.7. Because the CI does not cross 1.0 (which would indicate *no* increased odds of the outcome), this elevation in odds is statistically significant.

Number Needed to Treat (NNT)

This is the number of patients who need to be treated to produce one good outcome or prevent one additional bad outcome.

For example: Does a two-day treatment of intranasal ketamine reduce next-day MADRS scores to a greater extent than placebo? Eighteen people with depression received intranasal ketamine. As a control, the same 18 subjects were given placebo on a different date (ie, a "crossover" design). At 24 hours after ketamine, eight (44%) of the subjects had a response (a >50% reduction in MADRS score); after placebo, only one (6%) had a response.

$$NNT = \frac{1}{\text{absolute risk reduction (ARR)}}$$

$$ARR = \text{difference in response rates} = 0.44 - 0.06 = 0.38$$

$$NNT = \frac{1}{0.38} = 3 \quad \text{(NNTs are always given as whole numbers, and are rounded up)}$$

This means that three people need to be treated in order to see one response due to the drug (Lapidus KAB et al, *Biol Psychiatry* 2014;online ahead of print).

In clinical trials, often only the differences in outcome scores are reported (in this study, a difference in MADRS score of 7.6 points, which was statistically significant). However, the NNT helps us to identify how many patients would need to be treated with ketamine for one patient to respond.

Low NNTs are better (an ideal drug has NNT=1). And even though the FDA does not use NNT to approve drugs, all FDA-approved medications for bipolar disorder, for example, have NNT<10 (Ketter TA et al, *Acta Psychiatr Scand* 2011;123(3):175–189).

Keep in mind that an NNT=10 means that one patient will respond while nine people will not improve as a result of the drug. However, these nine may benefit from a placebo effect or simply through the natural evolution of symptoms. Using the drug is still acceptable if the risk and/or cost of the intervention is relatively small (see Citrome L, *Curr Psychiatry* 2008;6(3):66-71).

Effect Size

This is a measure of the size of the difference between two groups. Most commonly found in meta-analyses.

For example: How effective are antidepressants for chronic low back pain (CLBP)? Nine RCTs were evaluated, with a total of 504 patients. Measurements of pain severity were standardized across all studies, and rates of improvement were measured for antidepressants and for placebo.

Effect size can be calculated in several ways. The most common is known as Cohen's d (or simply d), also called a standardized mean difference (SMD).

The numerator is the difference in outcome scores (eg, a measure of pain or depressive symptoms), while the denominator is simply the standard deviation (SD) of these scores. If scores vary greatly, SD will be high and SMD will be low.

$$\text{SMD} = \frac{(\text{mean outcome on drug}) - (\text{mean outcome on placebo})}{\text{standard deviation}}$$

When these numbers were calculated for the nine antidepressant trials in CLBP, and then averaged, the result was 0.41 (Salerno SN et al, *Arch Int Med* 2002;162(1):19–24).

A higher SMD (or d) means a more effective intervention. When Cohen first developed criteria in 1988, he arbitrarily defined small, medium, and large effect sizes as 0.20, 0.50, and 0.80, respectively.

Incidentally, the effect size for antidepressants in depression has been estimated to be approximately 0.31 (Kirsch I et al, *PLoS Med* 2008;5(2):e45). Of course, effect sizes vary according to the severity of illness in subjects, the quality of the studies included, and many other factors.

Source: *TCPR,* July/August 2014, Vol 12, Issue 7&8, Research in Psychiatry