

Graph interpretation for the Applied Knowledge Test

Systematic reviews and meta-analyses can help busy clinicians draw together, evaluate and interpret large volumes of information drawn together from individual trials. They are placed at the summit of the hierarchy of research design, in part because of the precision and power that they provide. Understanding and interpreting the results of a systematic review or meta-analysis are therefore essential skills for any practicing clinician. In this article, we will review three of the common types of figure you may be presented with in a systematic review, a meta-analysis or in your Applied Knowledge Test (AKT). Caution - all three use fictional data!

Forest plots

A forest plot gives a graphical summary of the results of a meta-analysis [Sedgwick, 2015]. Imagine we wanted to assimilate the findings across all trials that assessed the relative risk of stroke for patients taking direct oral anticoagulants (DOACs) for primary stroke prevention in atrial fibrillation (figure 1). The first author and year are presented down the left column to identify each study that has been included. The reported relative risk from each individual study is represented by a green square box to the right. The bigger the box, the more weight the results of the study are contributing to the overall pooled result, usually because the study has more participants and therefore more precise outcomes. The black line either side of the green box represent the 95% confidence intervals. The wider the confidence interval the less certainty as to where the true relative risk lies within that range. To the far right the relative risk and 95% confidence interval for each study are given numerically.

In our example, the studies have been presented in subgroups of trials that have studied three new DOACs, 'Fixaban', 'Lobidatran' and 'Pravis'. Presenting the results in this way allows the reader to compare the pooled relative risk of each drug, represented by the summary blue diamond at the bottom of each subgroup. The overall combined relative risk for the three drugs compared to warfarin is given by the blue diamond at the bottom and the dotted red line that passes through it. The solid vertical line represents the line of no effect (relative risk = 0).

This forest plot shows that these fictional DOACs are associated with a 9% relative risk reduction for stroke compared to warfarin. The narrow confidence interval (-0.09 to -0.08) indicates that this is a very precise result. Fixaban appears to be associated with the greatest stroke reduction (RR -0.13, 95%CI -0.14, -0.12) and Lobidatran the least (RR -0.04, 95%CI -0.05, -0.03).

The I-squared score provides us with a measure of the amount of variation that is due to heterogeneity between the studies rather than due to chance alone. Causes of this might include differences in study design (e.g the dose of drugs given), length of follow-up, study inclusion criteria or quality of study. Generally speaking, I-squared values of 25%, 50% and 75% are respectively taken to mean low, moderate and high levels of heterogeneity. The I-squared score of 94.7% in our example therefore suggests there is a high degree of heterogeneity. We can also see this graphically, as results such as 'Lepay 2015' have 95%

confidence intervals that do not overlap with the other studies. Perhaps the Lepay study used a higher dose of Fixaban? When heterogeneity is high the authors should look at their results, explore the reasons for this heterogeneity and consider whether the original studies were actually too different to pool together in a meaningful way.

Box plots

Box plots are a helpful way of graphically summarising the distribution of a dataset [Driscoll et al., 2000]. The box itself shows the upper and lower quartile of the spread of data and the line through the middle of the box gives the mean result for that group. The 'whiskers' – lines that spread from the top and bottom of the box – show the overall spread of data for the group (the highest and lowest values).

For this example we use data comparing three imagined new ACE inhibitors, Canopril, Tadalapril and Zopril to placebo groups for blood pressure control (Figure 2). The mean BP is clearly lower in each ACEi group compared to placebo. In the Zopril group, some patients still had a systolic blood pressure >160mmHg and the mean, upper and lower quartile systolic BP results are all >140mmHg. Whilst all three drugs appear superior to placebo, these results suggest Tadalapril and Canopril are more effective than Zopril at lowering systolic blood pressure.

Funnel plots

A meta-analysis involves undertaking a systematic review of literature to find relevant research, but it is well known that studies that report positive findings are more likely to be published – so called 'publication bias' [Peters et al, 2010]. If a meta-analysis includes only published papers, we might expect the summary result to be biased towards a more positive outcome compared to when negative findings from unpublished sources are also included. Understanding this is important to help us interpret meta-analysis results and decide if we believe they are reliable.

Funnel plots present a scatter plot of each individual study within a meta-analysis, and are one way to assess for publication bias [Sedgwick, 2013]. Conventionally, they chart the outcome measure (e.g. relative risk, odds ratio or similar) on the x-axis against a degree of confidence in that result (e.g. standard error or standard deviation) on the y-axis. Standard error is a statistical term that measures the accuracy with which a sample represents a population. The y-axis is inverted, so that studies with a smaller standard error are at the top. The theory is that the smaller a standard error the more precise the result is likely to be. In the absence of publication bias, we expect results to scatter evenly across both sides so that the plot resembles a symmetrical (inverted) funnel. There will be a wider scattering of results across the x-axis at the bottom and more precise results with smaller standard errors clustering more closely around our overall estimate of effect size at the top of the y-axis.

In a funnel plot showing publication bias, as the standard error increases, studies tend to report more positive results. This will be seen with a clustering of results with a high degree of uncertainty over the precision of the outcome (i.e. a high standard error) towards the

positive end of the x-axis, with relatively few studies towards the negative end of the x-axis (Figure 3). Studies with lower standard errors are likely to continue to cluster around the overall effect size as larger results are more likely to be published regardless of outcome and also tend to have more precise outcome measures.

References and further information

Driscoll P, Lecky F and Crosby M (2000). An introduction to everyday statistics—2. *Emergency Medicine Journal*, 17(4):274-281. doi: [10.1136/emj.17.4.274](https://doi.org/10.1136/emj.17.4.274)

Peters JL, Sutton AJ, Jones DR et al (2010). Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *Statistics in Society*, 173(3): 575-591. Doi: 10.1111/j.1467-985X.2009.00629.x

Sedgwick P (2013). Meta-analyses: how to read a funnel plot. *British Medical Journal*, 346:f1342. Doi: 10.1136/bmj.f1342

Sedgwick P (2015). How to read a forest plot in a meta-analysis. *British Medical Journal*, 351:h4028. doi: 10.1136/bmj.h4028