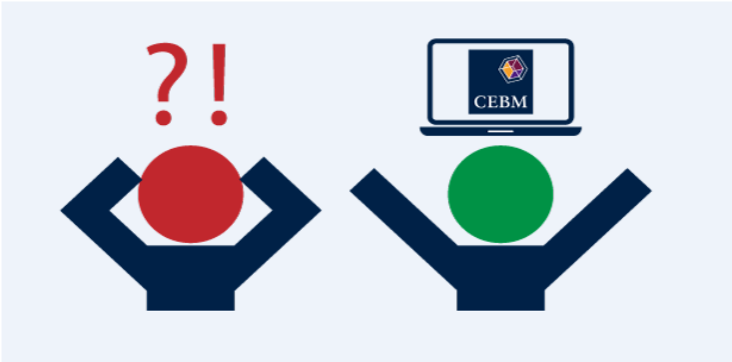
Tip for data extraction for meta-analysis – D1



What can I do when prognostic studies report measures of risk on different scales?

Kathy Taylor

Previously, I gave a set of tips for extracting data from diagnostic accuracy studies. I'll now look at a different study design, prognostic studies, and consider a problem with extracting hazard ratios, relative risks and odds ratios. I'll focus on hazard ratios, but my tip will also apply to relative risks and odds ratios.

First some background. Hazard ratios (HRs), also known as relative hazards, measure time-to-event data such as the time to a cardiovascular event or the onset of diabetes. A HR may be used to compare the risk of two groups. For example, a study reports that for people with type 1 diabetes and diabetic kidney disease, those with severe diabetic retinopathy have 46% higher risk of cardiovascular events compared to those without severe diabetic retinopathy (HR 1.46, 95% CI 1.11 to 1.92). A HR may also be used to express the change in risk associated with a specified change in a predictor variable. For example, another study reports that a 14mmHg increase an increase in the pre-awakening morning surge in systolic blood pressure in untreated hypertensive patients increases the risk of cardiovascular events by 33% (adjusted HR 1.33, 95% CI 1.01 to 1.06).

A HR comparing the risk of two groups from my first example could be pooled with HRs from similar studies that report the risk of the same outcome in the same two groups. However, a similar study to that in my second example may report a HR for the same outcome and same predictor but not the same change of in the predictor. These HRs need to be rescaled to a common change.

A bit of maths (see below if you're interested) shows us:

$$HR_y = (HR_x)^{\frac{y}{x}}$$

This equation shows the HR for an increase in y units of the predictor variable $\left(HR_y\right)$ is equal to the HR for an increase in x units (HR_x) raised to the power of $\frac{y}{x}$.

Also,

$$Lower\ limit\ of\ 95\%CI_y = lowerCI_y = (lowerCI_x)^{\frac{y}{x}}$$

$$Upper\ limit\ of\ 95\%\ CI_y = upperCI_y = (upperCI_x)^{\frac{y}{x}}$$

Let me show you a couple of examples from a review that I worked on. One study reported a 17% increased risk of cardiovascular mortality associated with a 5mmHg increase in 24-hour systolic blood pressure variability (HR 1.17, 95% CI 0.64 to 2.13) and another study reported a 3% increase in risk (HR 1.03, 0.93 to 1.13) per SD of the same predictor variable (15.6mmHg). I will rescale the second HR to a HR for a 5mmHg increase in blood pressure variability.

Taking x=15.6 and y=5

$$HR_5 = (HR_{15.6})^{\frac{5}{15.6}} = (1.03)^{\frac{5}{15.6}} = 1.01$$

$$Lower\ limit\ of\ 95\%\ CI_5 = lowerCI_5 = (lowerCI_{15.6})^{\frac{5}{15.6}} = (0.93)^{\frac{5}{15.6}} = 0.98$$

$$Upper\ limit\ of\ 95\%\ CI_5 = upperCI_5 = (upperCI_{15.6})^{\frac{5}{15.6}} = (1.13)^{\frac{5}{15.6}} = 1.04$$

i.e. HR 1.03 (0.93 to 1.13) for a 15.6mmHg increase rescales to 1.01 (0.98 to 1.04) for a 5mmHg increase in systolic blood pressure variability.

Here's a tip….

Hazard ratios, relative risks and odds ratios for different changes in a predictor variable can be rescaled to a common change

In my next post I'll look at beta coefficients.

Where did the equations come from?

(You can skip this if you are only interested in carrying out the calculations)

Hazard ratios (HRs) are estimated from Cox proportional hazards models (also known as Cox regression models). For these models,

$$\text{Expected hazard at time t} = h(t) = h(t)_o\, e^{(\beta_a P_a + \beta_b P_b + \beta_c P_c \ldots\ldots)}$$

Where P_a, P_b, P_c …. are predictor variables, $\beta_a, \beta_b, \beta_c$ …. are the coefficients, and $h(t)_o$ is the baseline hazard.

Consider increases in P_a only with all other predictors kept constant.

$$\text{HR for 1 unit increase} \quad = HR_1 = \frac{h(t)_0 e^{\beta_a(P_a+1)+\beta_b P_b+\beta_c P_c \ldots\ldots\ldots}}{h(t)_0 e^{\beta_a P_a+\beta_b P_b+\beta_c P_c \ldots\ldots\ldots}} = e^{\beta_a}$$

$$\text{HR for 5 units increase} \quad = HR_5 = \frac{h(t)_0 e^{\beta_a(P_a+5)+\beta_b P_b+\beta_c P_c \ldots\ldots\ldots}}{h(t)_0 e^{\beta_a P_a+\beta_b P_b+\beta_c P_c \ldots\ldots\ldots}} = e^{5\beta_a}$$

Terms cancel out because

$$e^{x+y} = e^x e^y$$

Assuming a constant hazard for each unit increase (this is the proportional hazards assumption of the Cox regression model):

$$\text{HR for x units increase} \quad = HR_x = e^{x\beta_a} = (HR_1)^x \Rightarrow HR_1 = (HR_x)^{\frac{1}{x}}$$

Similarly for a HR for an increase of y units and substituting

$$\text{HR for y units increase} \quad = HR_y = e^{y\beta_a} = (HR_1)^y = (HR_x)^{\frac{y}{x}}$$

The same calculations apply to the upper and lower confidence limits.

Therefore, scaling from x units to y units increase in the predictor variable

$$HR_x(lowerCI_x \ to \ upperCI_x) \ rescales \ to \ (HR_x)^{\frac{y}{x}} \left\{(lowerCI_x)^{\frac{y}{x}} \ to (upperCI_x)^{\frac{y}{x}}\right\}$$

Dr Kathy Taylor teaches data extraction in Meta-analysis. This is a short course that is also available as part of our MSc in Evidence-Based Health Care, MSc in EBHC Medical Statistics, and MSc in EBHC Systematic Reviews.

Follow updates on this blog, related news, and to find out about other examples of statistics being made more broadly accessible on Twitter @dataextips