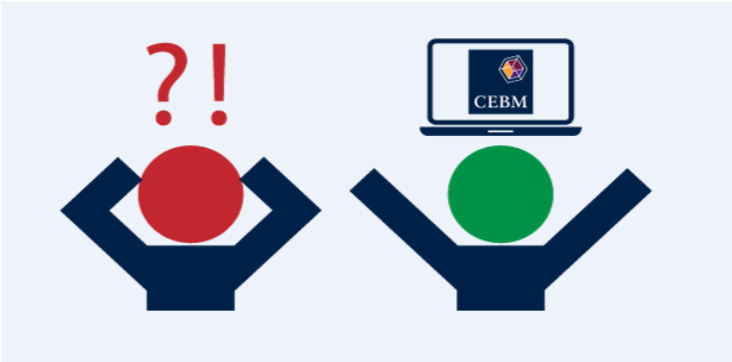
Tip for data extraction in meta-analysis - 3



What if the sensitivity or specificity is also not reported?

Kathy Taylor

In the previous two posts I considered data extraction for diagnostic accuracy studies. First I showed how to derive the 2x2 diagnostic accuracy table from the reported sensitivity, specificity, prevalence and study size. Then I showed how, if the prevalence is also not reported, it could be calculated using the number with the disease, the positive predictive value (PPV) or the negative predictive value (NPV). In this post I'll explain how to deal with the less common case when a sensitivity or specificity is not reported.

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

$$Sensitivity = \frac{PPV \ \times (1 - Specificity) \ \times (1 - Prevalence)}{Prevalence \times (1 - PPV)}$$

$$Specificity = \frac{NPV \ \times (1 - Sensitivity) \ \times Prevalence}{(1 - Prevalence) \ \times (1 - NPV)}$$

Be careful – statistics such as prevalence can be reported as percentages or as decimal fractions. We're assuming that sensitivity, specificity and prevalence are reported as fractions.

Let me show you an example. In a review that I'm currently working on, most studies report the sensitivity and specificity and some also report the PPV and NPV, but one study only reports the sensitivity and NPV for the reported subgroups. It's a study of the detection of left ventricular systolic dysfunction using brain natriuretic peptide (BNP). For the high-risk subgroup, using BNP measured with RIK 9076 kit from Peninsular Laboratories at a threshold 8 pg/mL and compared to echocardiography, the sensitivity is reported at 94% and the NPV at 99%. The size of this subgroup is 269 and 16 are reported as having left ventricular systolic dysfunction.

Using the first equation I showed in my previous post (<link>EBHC KT blog2</link>)

$$Prevalence = \frac{Number\ with\ the\ disease\ in\ the\ subgroup}{Total\ in\ the\ subgroup} = \frac{16}{269} = 0.059\ or\ 5.9\%$$

This study reports percentage inputs, so we first need to convert these to decimal fraction inputs by dividing the percentages by 100 (e.g. NPV of 99% becomes 0.99). Then applying the equation from above and using the NPV

$$Specificity = \frac{0.99\ \times (1 - 0.94)\ \times 0.059}{(1 - 0.059)\ \times (1 - 0.99)} = \frac{0.00351}{0.00941} = 0.37\ or\ 37\%$$

Here's a tip....

The sensitivity and specificity can be calculated from other statistics

My next set of tips will focus on some common data extraction problems in prognostic studies. This will involve data being log-transformed so, in preparation, I'll be posting some notes next week on logarithms and log-transformations.

Where did the equations come from?

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

The equations are those given previously (<link>EBHC KT blog 2</link>) and rearranged.

Rearranging

$$Prevalence = \frac{PPV\ \times (1 - Specificity)}{Sensitivity\ \times (1 - PPV) + PPV \times (1 - Specificity)}$$

becomes

$$Sensitivity = \frac{PPV\ \times (1 - Specificity)\ \times (1 - Prevalence)}{Prevalence \times (1 - PPV)}$$

Rearranging

$$Prevalence = \frac{Specificity\ \times (1 - NPV)}{NPV\ \times (1 - Sensitivity) + Specificity \times (1 - NPV)}$$

becomes

$$Specificity = \frac{NPV\ \times (1 - Sensitivity)\ \times Prevalence}{(1 - Prevalence)\ \times (1 - NPV)}$$

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