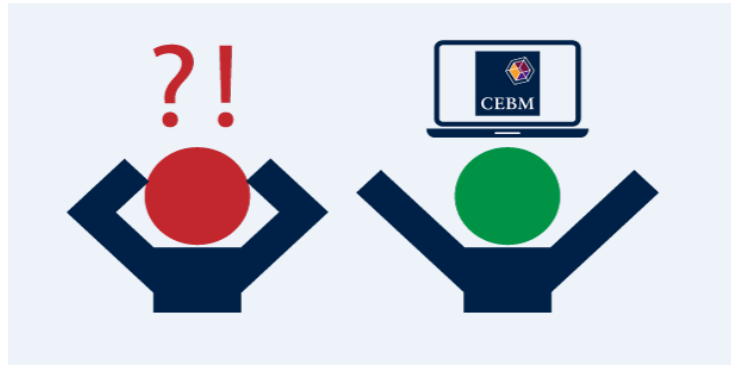


Tip for data extraction in meta-analysis - 20



What if neither sensitivity nor specificity are reported?

Kathy Taylor

In this post I return briefly to data extraction problems in diagnostic accuracy studies, as there's another situation that I've not covered.

Previously I showed how to derive the 2x2 diagnostic accuracy table from the reported sensitivity, specificity, prevalence and study size (total). Then I showed how, if the prevalence is also not reported, it could be calculated from the number with the disease and the positive predictive value (PPV) or the negative predictive value (NPV). Then I looked at the less common case where either the sensitivity or the specificity is not reported and I showed how this problem can be solved using the prevalence, NPV and PPV. In this post, I'll look at a rarer case where both sensitivity and specificity are not reported.

To summarise, I've shown how it's possible to derive a 2x2 diagnostic accuracy table when the following statistics are reported:

- Total, prevalence, sensitivity and specificity
- Total, number with the disease, sensitivity and specificity
- Total, NPV, sensitivity and specificity
- Total, PPV, sensitivity and specificity
- Total, prevalence, sensitivity and PPV
- Total, prevalence, specificity and NPV
- Total, number with the disease, sensitivity and PPV
- Total, number with the disease, specificity and NPV

In this post, to the above list, I'm going to add the following

Total, prevalence, PPV and NPV

Total, number with the disease, PPV and NPV

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

$$\text{Sensitivity} = \frac{PPV}{Prevalence} \times \left(\frac{1 - Prevalence - NPV}{1 - PPV - NPV} \right)$$

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

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

Let me show you an example. Consider 180 patients of which 100 have a disease and a test correctly identifies 20 patients with the disease and correctly produces a negative test result for 40 of the patients without the disease.

	Disease	No disease	
Test positive	80	20	100
Test negative	20	60	80
	100	80	180

Prevalence=100/180=5/9
 PPV=80/100=4/5
 NPV=60/80=3/4
 Sensitivity=80/100=4/5
 Specificity=60/80=3/4

Using the equations

$$\text{Sensitivity} = \frac{\frac{4}{5} \left(1 - \frac{5}{9} - \frac{3}{4} \right)}{\frac{5}{9} \left(1 - \frac{4}{5} - \frac{3}{4} \right)} = \frac{36}{45} = \frac{4}{5}$$

$$\text{Specificity} = \frac{\frac{3}{4} \left(\frac{5}{9} - \frac{4}{5} \right)}{\left(1 - \frac{5}{9} \right) \left(1 - \frac{3}{4} - \frac{4}{5} \right)} = \frac{45}{60} = \frac{3}{4}$$

Here's another example using data from a published [study](#). It's based on the detection of left ventricular systolic dysfunction using N-terminal atrial natriuretic peptide. For severe left ventricular dysfunction (n=8) at threshold 1.4 ng/mL the NPV was 100% and PPV was 22%. Of 134 patients, 8 had 8/134 or 6%. Using the equations above, the specificity is calculated as

$$\text{Sensitivity} = \frac{0.22 \left[\frac{1 - 0.06 - 1}{1 - 0.22 - 1} \right]}{0.06} = 1.0$$

$$\text{Specificity} = \frac{1.00 \times (0.06 - 0.22)}{(1 - 0.06)(1 - 0.22 - 1.00)} = 0.77$$

Here's a tip....

When constructing 2x2 diagnostic accuracy tables, if both the sensitivity and specificity are missing, you may be able to calculate them from other statistics.

My next set of tips will focus on some common and less common data extraction problems in intervention studies. I'll show you what you can do when you have a missing mean, standard deviations or sample size.

Where did the equations come from?

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

Previously I gave the following equations

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

(equation 1)

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

(equation 2)

Substituting equation 2 into equation 1 and rearranging for Sensitivity gives

$$\text{Sensitivity} = \frac{PPV}{\text{Prevalence}} \times \left(\frac{1 - \text{Prevalence} - NPV}{1 - PPV - NPV} \right)$$

Substituting equation 1 into equation 2 and rearranging for Specificity gives

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

The above equations use Prevalence. I [previously](#) showed how to calculate prevalence from the total number of patients and number with the disease.

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](#)