DIAGNOSTIC ACCURACY STUDIES

Are the results of the study valid?

Was the diagnostic test evaluated in a Representative spectrum of patients (like those in whom it would be used in practice)?

What is best?

It is ideal if the diagnostic test is applied to the full spectrum of patients - those with mild, severe, early and late cases of the target disorder. It is also best if the patients are randomly selected or consecutive admissions so that selection bias is minimized.

Where do I find the information?

The Methods section should tell you how patients were enrolled and whether they were randomly selected or consecutive admissions. It should also tell you where patients came from and whether they are likely to be representative of the patients in whom the test is to be used.

In this paper

Yes	No	Unclear
☐	☐	☐

Comment:

Was the reference standard applied regardless of the index test result?

What is best?

Ideally both the index test and the reference standard should be carried out on all patients in the study. In some situations where the reference standard is invasive or expensive there may be reservations about subjecting patients with a negative index test result (and thus a low probability of disease) to the reference standard. An alternative reference standard is to follow-up people for an appropriate period of time (dependent on disease in question) to see if they are truly negative.

Where do I find the information?

The Methods section should indicate whether or not the reference standard was applied to all patients or if an alternative reference standard (eg follow-up) was applied to those who tested negative on the index test.

In this paper

Yes	No	Unclear
☐	☐	☐

Comment:

Was there an independent, blind comparison between the index test and an appropriate reference ('gold') standard of diagnosis?

What is best?

There are two issues here. First, the reference standard should be appropriate - as close to the 'truth' as possible. Sometimes there may not be a single reference test that is suitable and a combination of tests may be used to indicate the presence of disease.

Second, the reference standard and the index test being assessed should be applied to each patient independently and blindly. Those who interpreted the results of one test should not be aware of the results of the other test.

Where do I find the information?

The Methods section should have a description of the reference standard used and if you are unsure of whether or not this is an appropriate reference standard you may need to do some background searching in the area.

The Methods section should also describe who conducted the two tests and whether each was conducted independently and blinded to the results of the other.

In this paper

Yes	No	Unclear
☐	☐	☐

Comment:

What were the results?

Are test characteristics presented?

There are two types of results commonly reported in diagnostic test studies. One concerns the accuracy of the test and is reflected in the sensitivity and specificity. The other concerns how the test performs in the population being tested and is reflected in predictive values (also called post-test probabilities). To explore the meaning of these terms, consider a study in which 1000 elderly people with suspected dementia undergo an index test and a reference standard. The prevalence of dementia in this group is 25%. 240 people tested positive on both the index test and the reference standard and 600 people tested negative on both tests. The first step is to draw a 2 x 2 table as shown below. We are told that the prevalence of dementia is 25% therefore we can fill in the last row of totals - 25% of 1000 people is 250 - so 250 people will have dementia and 750 will be free of dementia. We also know the number of people testing positive and negative on both tests and so we can fill in two more cells of the table.

Reference standard

		+ve	-ve	
Index test	+ve	240		
	-ve		600	
		250	750	1000

By subtraction we can easily complete the table:

Reference standard

		+ve	−ve	
Index test	+ve	240	150	390
	−ve	10	600	620
		250	750	1000

Now we are ready to calculate the various measures.

What is the measure?	What does it mean?
Sensitivity (Sn) = the proportion of people with the condition who have a positive test result.	The sensitivity tells us how well the test identifies people with the condition. A highly sensitive test will not miss many people.
In our example the Sn = 240/250 = 0.96	10 people (4%) with dementia were falsely identified as not having it. This means the test is fairly good at identifying people with the condition.
Specificity (Sp) = the proportion of people without the condition who have a negative test result.	The specificity tells us how well the test identifies people without the condition. A highly specific test will not falsely identify many people as having the condition.
In our example the Sp = 600/750 = 0.80	150 people (20%) without dementia were falsely identified as having it. This means the test is only moderately good at identifying people without the condition.
Positive Predictive Value (PPV) = the proportion of people with a positive test who have the condition.	This measure tells us how well the test performs in this population. It is dependent on the accuracy of the test (primarily specificity) and the prevalence of the condition.
In our example the PPV = 240/390 = 0.62	Of the 390 people who had a positive test result, 62% will actually have dementia.

Negative Predictive Value (NPV) = the proportion of people with a negative test who do not have the condition.

In our example the NPV = 600/610 = 0.98

This measure tells us how well the test performs in this population. It is dependent on the accuracy of the test and the prevalence of the condition.

Of the 610 people with a -ve test , 98% will not have dementia.

Applicability of the results

Were the methods for performing the test described in sufficient detail to permit replication?

What is best?

The article should have sufficient description of the test to allow its replication and also interpretation of the results.

Where do I find the information?

The Methods section should describe the test in detail.

In this paper

Yes	No	Unclear
☐	☐	☐

Comment:

DIAGNOSTIC ACCURACY STUDIES

CEBM

UNIVERSITY OF OXFORD