Harvard Pilgrim Health Care Institute

CLINICAL EPIDEMIOLOGY AND POPULATION HEALTH

Key Points – Diagnostic Testing and Screening

To use diagnostic tests correctly, understanding the following concepts is critical.

Sensitivity

- Among patients with the disease, how often is the test correct?
  - = Correct tests among patients who have disease
  - = True positives ÷ (true positives + false negatives)
- A perfectly sensitive test never misses disease.
- Vertical in 2x2 table

Specificity

- Among patients without the disease, how often is the test correct?
  - = Correct tests among patients who do not have the disease
  - = True negatives ÷ (true negatives + false positives)
- A perfectly specific test never classifies a well person as sick.
- Vertical in 2x2 table

Pre-test probability *(i.e., pre-test probability of disease)*

- The chance that the patient has the disease before you have the test results. This is simply the *prevalence* of disease among patients with the same characteristics.
- Determined by a combination of descriptive epidemiology and clinical gestalt
- Horizontal in 2x2 table

Post-test probability *(i.e. post-test probability of disease)*

- The chance that the patient has the disease after you have the test results.
- This is also called the ***predictive value*** of the test.
- It depends on the test's sensitivity & specificity, and on the pre-test probability of disease.
- Horizontal in 2x2 table

Threshold probability

- The probability above or below which you will take action (e.g., order an additional test, prescribe a treatment).
- Varies with the clinical condition and the potential benefits of treatment and harms of undertreating, overtreating, or doing further testing.

Cutoff

- We often talk about diagnostic tests as "positive or negative." Even in these cases, there is often an underlying value of a continuous variable that has been chosen as the cutoff for "positive" vs. "negative." Choosing an appropriate cutoff depends on how we weigh the risks of false positive and false negative errors.

Diagnostic test results are often represented by 2x2 tables. Below is an example, with a rapid antigen test (from a throat swab) as the test and streptococcal pharyngitis (strep throat) confirmed by culture as the disease:

		Streptococcal Pharyngitis (the DISEASE)	
		PRESENT	ABSENT
Rapid antigen test (the TEST)	POSITIVE	True Positive (TP)	False Positive (FP)
	NEGATIVE	False Negative (FN)	True Negative (TN)

Sensitivity: among those with the disease (TP + FN), how many have a positive test result (TP)?
= TP/(TP + FN)

Specificity: among those without the disease (TN + FP), how many have a negative test result (TN)?
= TN/(TN + FP)

Sensitivity and specificity are attributes of a test. However, clinicians are more concerned about the probability of their patient having the disease in question- the *post-test probability of disease*.

Positive Predictive Value (PPV) = TP/(TP +FP)
Is the chance that a positive test result accurately signifies the presence of the disease (also called Predictive Value Positive or Post-test Probability *of disease* given a positive test).

Negative Predictive Value (NPV) = TN/(TN +FN)
Is the chance that a negative test result accurately signifies the absence of the disease (it is the Post-test Probability of *not having disease* given a negative test).

Note that the Post-test Probability *of disease* given a negative test = (1-NPV). This is the chance that a patient has the disease after a negative test result.

All post-test probabilities are dependent on the test characteristics, but also on the pre-test probability of disease. For example, a positive test result (even for a test with reasonable sensitivity and specificity) in a patient with a very low pre-test probability may still not indicate a high probability of a disease. This is critical in interpreting diagnostic test information.

Determining the probability of disease by knowing the pre-test probability and the attributes of the test is an application of *Bayes Theorem,* which describes the probability of an event based on prior knowledge of conditions that might be related to the event.

Likelihood Ratios and Odds
Likelihood ratios are the probability of a given result in a person with the disease divided by the probability of that result in people without the disease. They give a measure of the test's ability to discriminate between people with and without the disease.

$$LR^{+} = \frac{sensitivity}{(1 - specificity)} = \frac{TruePositiveRate}{FalsePositiveRate}$$

$$LR^{-} = \frac{(1 - sensitivity)}{specificity} = \frac{FalseNegativeRate}{TrueNegativeRate}$$

Multiplying the pre-test odds of a disease by the LR results in post-test odds. The formulas for converting probability to odds and odds back to probability are:

$$odds = \frac{p}{1-p} \qquad p = \frac{odds}{1+odds}$$

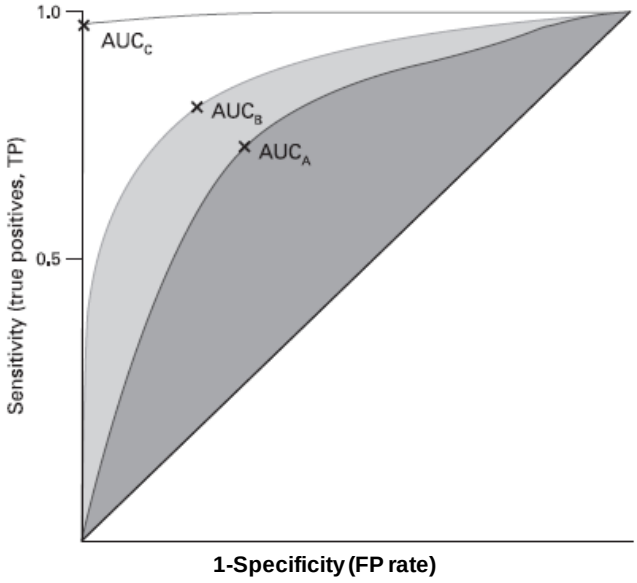
Likelihood ratios combine sensitivity and specificity, are independent of the prevalence of disease, facilitate combining tests, and allow comparisons of information from different cutoff levels of the same test.

As with anything, estimates from studies of sensitivity, specificity, and LRs can be affected by chance (confidence intervals can be calculated) or bias.

Receiver Operating Characteristic (ROC) Curves

Diagnosis, or assessment of the probability of a disease being present, can also be based on continuous markers. A perfect test would completely separate normal from abnormal, but such tests generally do not exist. For tests with continuous results, *we* choose the cut-off below which we will consider normal and above which we will consider abnormal (or vice versa). In doing so, there is always a trade-off. If we change our cutoff to increase sensitivity, then specificity goes down. If we change our cutoff to increase specificity, then sensitivity goes down. Such trade-offs are easily shown in Receiver Operating Characteristic (ROC) Curves. An ROC curve is a plot of the true positive rate (sensitivity) against the false positive rate (1-specificity; recall that specificity reflects true negatives) for the possible cutpoints of a diagnostic test. The closer the curve follows the left-hand border and the top border of the ROC space (passes through the upper left hand corner), the more accurate the test. Such a test has both high sensitivity and high specificity. The closer the curve comes to the 45-degree diagonal line the less accurate the test.

We can also compare different tests, or different threshold values on the same test, by graphing their ROCs together and showing their relative areas-under-the-curve (AUCs) (see below, which might compare tests A vs. B vs. C, or compare threshold A vs. B vs. C on the same test).



Introduction to Screening

Screening is the use of a test to identify a disease before any clinical signs or symptoms manifest. Screening methods may include asking questions, physical examination (e.g., skin exam for atypical moles), blood tests (e.g., cholesterol or PSA level), or imaging studies. Typically, prevalence of the condition is lower than in clinical diagnosis situations, in which patients present with symptoms. There are some general considerations regarding which conditions make sense to screen for, e.g., whether the condition is detectable before symptoms emerge, whether early treatment confers benefit, and whether benefits to those with a positive screening test result outweigh the harms of screening (including the harms of false positive results).

Assessment of a screening test or strategy can be subject to particular ***biases***:

Lead time bias

- People with the disease who get screened and then get treatment survive longer than people with the disease who do not get screened. You may mistakenly conclude that screening (followed by treatment) cause improved survival, when what is really going on is that patients diagnosed earlier have more time to survive, even if treatment isn't helpful and their actual life expectancy may be no different (or, perhaps even worse).

Length bias

- Cases of a particular condition (e.g., cancer) that are progressing more slowly will be more prevalent in a screened population than those that are rapidly progressing. Remember, you can only call it "screening" if you don't already know the patient has the disease. So if some patients with the disease spend more time "having it without being sick from it," they are the ones you'll find by screening. In contrast, patients with rapid disease progression spend less time "having it without being sick from it." That means that patients with positive screening tests appear to have better outcomes, but it is not because screening+treatment is better – instead it's because the disease in patients detected by screening is different from that in patients who present with symptoms.