ERIC Notebook

March 2000 Issue 11

UNC SCHOOL OF PUBLIC HEALTH DEPARTMENT OF EPIDEMIOLOGY

Michel Ibrahim, MD, PhD
Director of Education Program

Lorraine Alexander, DrPH

Carl Shy, MD, DrPH

Sherry Farr, Graduate Research Assistant

VA ERIC EPIDEMIOLOGIC RESEARCH & INFORMATION CENTER Durham, North Carolina

Ron Horner, PhD
ERIC Director

http://hsrd.durham.med.va.gov/ERIC/

The ERIC Notebook is funded by the Department of Veterans Affairs (DVA), Veterans Health Administration (VHA), Cooperative Studies Program (CSP), to promote the strategic growth of the epidemiologic capacity of the DVA.

Assessment of Diagnostic and Screening Tests

Diagnostic and screening tests are used to detect the presence or severity of disease in individuals. Clinicians rely on these tests to make decisions in treating patients. Therefore, the performance of a new test must be assessed before being used in a clinical setting. The performance of any new diagnostic or screening test is assessed by comparing actual test results to the patient's true disease status (as assessed by a gold standard). The four measures used to evaluate a new test are the sensitivity, specificity, and positive and negative predictive values.

The gold standard

"Gold standard" is a term for the most definitive diagnostic procedure, e.g. microscopic examination of a tissue specimen, or the best available laboratory test, e.g. serum antibodies to HIV. Sometimes it can refer to a comprehensive clinical evaluation, e.g. clinical assessment of arthritis. "Gold standard" procedures can often be costly, invasive and/or uncomfortable. New tests that are less invasive and less expensive are compared against them in these terms. A test which detects a marker in the blood for prostate cancer may not be as sensitive as taking a biopsy from the prostate itself, but the discomfort of biopsy may make the blood assay a better alternative.

Calculating the test results

A table like the one below is used to group individuals into one of four disease-test categories.

	Truth (Gold Standard)	
Results of test	Number with Disease Present	Number with Disease Absent
Positive Test	a = True Positive	b = False Positive
Negative Test	c = False Negative	d = True Negative

Prevalence of the disease =
$\frac{\text{True positives}}{\text{Total population}} = (a+c) / (a+b+c+d)$

Sensitivity vs. Specificity

Sensitivity and specificity are measures that assess the validity of diagnostic and screening tests. These measures reflect how well the test is detecting the disease and classifying individuals into disease and non-disease groups.

- Sensitivity (Se) describes how well the test detects disease in all who truly have disease, or the percent of diseased individuals who have positive test results.
- Specificity (Sp), on the other hand, describes how well the test is detecting non-diseased individuals as truly not having the disease, or the percent of non-diseased individuals who have negative test results.

$Se = \frac{\text{True-positives} \times 100}{\text{True-positives + False-negatives}} = a / (a+c) \times 100$

$Sp = \frac{\text{True-negatives} \times 100}{\text{True-negatives + False-positives}} = d / (b+d) \times 100$

A highly sensitive test means that a large percent of people who have disease are classified correctly as having the disease. A highly specific test means that a large percent of individuals without disease are classified correctly as not having disease. An ideal test would be both highly sensitive and highly specific, where disease would be detected in 100% of those who truly have disease (100% sensitivity), and disease would be ruled out in 100% of those who are truly disease-free (100% specificity).

- For example, if a test is 95% sensitive and 98% specific, then 5% of the diseased individuals will have negative test results (the test is incorrectly classifying 5% of the diseased individuals), and 2% of the disease-free individuals tested will have positive test results (the test is incorrectly classifying 2% of the disease-free individuals).

False-positives and false-negatives
A false positive is an individual who is incorrectly diagnosed as a case when, in fact, they do not have the disease. A false negative is an individual who is incorrectly diagnosed as a non-case, when in fact the person does have the disease.

- 100% - %sensitivity = % false negatives
- 100% - %specificity = % false positives

Positive and negative predictive values
The positive predictive value (PV+) is the percent of positive tests that are truly positive. The negative predictive value (PV-) is the percent of negative tests that are truly negative. Like sensitivity and specificity, PV+ and PV- also show how well the test is classifying individuals into disease and non-disease groups, but the denominator for PV+ is the total number of persons who test positive (a + b), while that for PV- is the total number who test negative (c + d). A test with a high PV+ value means that there is only a small percent of false-positives within all the individuals with positive test results. A test with a high PV- value means that there is only a small percent of false-negatives within all the individuals with negative test results.

Positive predictive value (PV+) =
True Positives / All positives on the test = a / (a+b)

Negative predictive value (PV-) =
True Negatives / All negatives on the test = d / (c+d)

- For example, a certain test, e.g. a stress ECG, which has a PV+ of 90% and a PV- of 95% is used to screen 5,000 people for coronary heart disease. Forty percent of the individuals (2,000 people) have positive test results and 60% (3,000 people) have negative test results. If the gold standard for CHD found that 1,800 of those who tested positive (90% of 2,000) truly have CHD, and 2,850 of those who tested negative (95% of 3,000) are truly non-cases.

Pros and cons of specificity and sensitivity
Ideally, an investigator would prefer a diagnostic test that is both 100% sensitive and 100% specific. However, this scenario rarely occurs. It is important in clinical decision-making to know the sensitivity and specificity of the test you are conducting and to weigh the pros and cons of using tests with different levels of sensitivity and specificity. For instance, if a disease is not life threatening if left untreated, the costs of treatment are high, and invasive surgery is required, then a very specific diagnostic test is preferred over a more sensitive test. If the disease under study is life threatening if left untreated, and the survival rate is improved with immediate treatment, then the sensitivity of a diagnostic test is of greater importance than its specificity.

The prevalence affects the test measures
The prevalence of a disease affects the PV+ and PV- values. If a disease has a low prevalence and the test being used to assess disease in individuals is not 100% sensitive or 100% specific, as will most likely be the case, then false-positives may overwhelm the positive test results.

- For example, schizophrenia has a low prevalence in the U.S. at around 1%. A new diagnostic test which is 99% sensitive and 99% specific is used to screen 10,000 patients for schizophrenia. Of those 10,000, we would expect 100 to truly be suffering from schizophrenia, or 1% of our population. Of those 100, 99 (99% of 100) would have positive test results. Of the 9,900 who are truly without disease, 9801 (99% of 9,900) would be classified as disease-free. However, there would be 99 (9,900-9801) false positives. This test would give 198 (99+99) positive test results. Therefore, even with a test that is 99% sensitive and 99% specific, the PV+ would only be 50% (99/198).

Self-Evaluation

Q1: The enzyme-lined immunosorbent assay (ELISA) is the common test used as a first screen for HIV antibodies in blood. Assume that the sensitivity of ELISA is 97.0% and the specificity is 99.8%. Assume that the total number of persons being tested for HIV is 100,000; therefore, a+b+c+d = 100,000. Assume that prevalence of HIV infection in this population is 40 per 100,000. Make a 2x2 table and calculate the values for a, b, c, and d. For example, because sensitivity = [a/(a + c)], it follows that a = (sensitivity) x (a + c). Adjust your calculated values for a, b, c, and d to the nearest whole number.

Q2: What is the positive predictive value (PV+) of ELISA in the population of 100,000?

Q3: Now apply the ELISA screen to a population group at high risk of AIDS, such as intravenous drug users. If you had access to 1000 IV drug users and could obtain permission to test them for HIV antibody status, what value will you obtain for the PV+ in this population? Use the same sensitivity (97.0%) and specificity (99.8%) values you used in Question one. Assume past studies have estimated that 20% of IV drug users are HIV positive.

Q4: Comparing your answers to Questions 1 and 3, you can conclude that the positive predictive value of a screening test is determined by what three factors?

Answers
1. (a + b + c + d) = 100,000 persons to be screened. a + c=40, (40 persons are HIV infected.) Sensitivity = 0.97=[a/(a + c)]
a=0.97(a + c)=0.97(40)=39 c=[(a + c) - a]=1

Specificity = 0.998 = [d/(b + d)]
(b + d)= Total-(a + c) = 99,960
d = 0.998 (b + d) = 99,760
b = (b + d)-d = 200
Therefore, the tabulated results are as follows:
Positive: a = 39; b = 200; a + b = 239; Negative: c = 1;
d = 99,760; c + d = 99,761; Total: a + c = 40; b + d = 99,960;
a + b +c + d = 100,000

2. Positive predictive value: among those who test positive, what fraction actually have HIV antibodies = a/(a + b) = 0.16. Only 16% of those who test positive actually have HIV antibodies.

3. Prevalence = 20% of 1000 = 200
a + c = 200 actually have HIV antibodies
a = Sensitivity x (a + c) = 0.97(200) = 194
c = 200-194 = 6
d = Specificity x (b + d) = 0.998(800) = 798
b = 800-798 = 2
PV+ = a/(a + b) = 194/196 = 0.990

4. PV+ is determined by the (1) sensitivity of the test, (2) the specificity of the test, and (3) prevalence of the disease.

Glossary

False positive - An individual who is incorrectly diagnosed as having disease.

False negative - An individual who is incorrectly diagnosed as not having disease.

Gold standard – The most definitive diagnostic procedure to detect disease.

Negative predictive value – The percent of negative tests that are truly negative.

Positive predictive value – The percent of positive tests that are truly positive.

Sensitivity – The percent of diseased individuals who have positive test results.

Specificity – The percent of non-diseased individuals who have negative test results.

Announcement

Short Courses on the Internet for Continuing Medical Education Credit in Basic Epidemiologic Methods

The Epidemiologic Research and Information Center (ERIC) at the Durham, North Carolina VA Medical Center and The Department of Epidemiology, School of Public Health, the University of North Carolina at Chapel Hill announce the offering of several short two hour courses in basic epidemiologic methods for Continuing Medical Education (CME) credit. These courses will be available over the Internet and will be offered free of charge for a period of three months. The topics that these courses will cover include: basic study designs such as case control, cohort and cross sectional studies, as well as selection bias, outbreak investigations and such epidemiologic measures as incidence and prevalence. These courses have been planned and implemented in accordance with the Essential Areas and policies of the Accreditation Council for Continuing Medical Education (ACCME) through the joint sponsorship of The School of Medicine and The School of Public Health of the University of North Carolina at Chapel Hill. The School of Medicine of The University of North Carolina at Chapel Hill is accredited by the ACCME to provide continuing medical education for physicians and takes responsibility for the content, quality and scientific integrity of this CME activity.

The first two courses, Cohort Studies and Outbreak Investigations will be available on the Internet beginning March 1, 2000. The opening dates of the other courses will be announced as they become available.

Courses for CME credit can be accessed by going to the following URL:
http://cdlhc.sph.unc.edu/courses/eric/
Please feel free to forward this announcement to anyone you think may be interested.

ERIC NOTEBOOK IS PRODUCED BY THE EDUCATIONAL ARM (MICHEL A. IBRAHIM, MD, PHD, DIRECTOR) OF THE EPIDEMIOLOGIC RESEARCH AND INFORMATION CENTER AT DURHAM, NC (RON HORNER, PHD, DIRECTOR)

If you would like to receive ERIC Notebook please fill out the form below:

Name: _____

Degree(s):_____

Address:_____

City, State, Zip:_____

Telephone Number:_____

Fax Number:_____

E-mail Address:_____

Please fax to: 919-416-5836 – Attn: Beth Armstrong **Or**

Mail to : Beth Armstrong, ERIC Program Manager, VA Medical Center (152), 508 Fulton Street, Durham, NC 27705

Upcoming Topics

Ecologic studies
Confounding

Please let Beth Armstrong know which topics are of special interest to you so that we can include them in a future issue.

BETH ARMSTRONG, ERIC PROGRAM MANAGER
VA MEDICAL CENTER (152)
508 FULTON STREET
DURHAM, NC 27705