Rational Use of Diagnostic Tests: From the Old Standards to the Latest in Genetic Testing

Michael G. Shlipak, MD, MPH
Professor of Medicine, UCSF
Chief, Division of General Internal Medicine
San Francisco VA Medical Center

Introductions

• Job Description?
  General Internal Medicine? Family Practice? Nurse Practitioner/Physician’s Assistant? Pharmacist? Other?

• What is your comfort level for diagnostic tests?
  0 Not at all 1-3 4-7 8-10 Very comfortable

• Are you comfortable with the concept of:
  Sensitivity?
  Specificity?
  Likelihood ratios?

• Have you used a likelihood ratio in clinical practice?

Goals of this Workshop

• To improve your confidence for interpreting test results
• To emphasize how disease likelihood (probability) influences the meaning of test results
• To advance comprehension of sensitivity and specificity, and their limitations
• To introduce the concept and usefulness of likelihood ratios
• To practice using Likelihood Ratios (with a Ruler!)

Do I Really Need to Understand Likelihood Ratios?

• Peterson LR, Robicsek A. Does my patient have Clostridium difficile infection? Ann Intern Med. 2009
• Marx A et al. Meta-analysis: accuracy of rapid tests for malaria in travelers returning from endemic areas. Ann Intern Med. 2005
Why Do We Obtain Tests?

• To detect disease
• To help us estimate the probability of disease
• Move from indecision to:
  ◦ Action (treatment)
  ◦ Inaction (reassurance)
  ◦ More testing
• “Tests” include the following:
  ◦ History
  ◦ Physical exam
  ◦ Radiology
  ◦ Lab tests

Can you Give an Example of a Test in Each Category?

• History
• Physical exam
• Radiology
• Lab tests

Let's take the example of patients who present with intermittent chest pain in your clinic.

Reassurance  |  Order Test  |  Treat/Percath/PCI

Test threshold  |  Treatment threshold

Probability of disease

0%  |  100%

Another Example:

• Lots of patients present with cough. Sometimes we order CXR. Sometimes we just prescribe Abx for outpatient PNA. Sometimes we just treat the symptoms of cough.

Possible Pneumonia in a Patient with Cough

0%  |  100%

Reassurance  |  CXR  |  Empiric Antibiotics

In clinical practice we rarely articulate the disease probability, but it affects our decision making everyday.
The best tests are definitive. A good test moves us across action thresholds. The best tests are definitive.

Likelihood of Disease Depends on 2 Things
1. Where you started from (low, medium, high risk): prior probability
2. Length and direction of the “arrow”: likelihood ratio (LR)

• Basic paradigm:
  - Prior probability + LR from test = post-test probability

Likelihood of Disease Depends on 2 Things

Example Using a Hypothetical Test

Order a Test
Reassurance
Treatment

HIV-
Negative test
0%

HIV+
Positive test
100%

“What We Thought Before” = Prior Probability

- Likelihood of a specific disease
- Estimated before knowing test result (0-100%)
- We should try to force ourselves to quantify probability.

How Do We Get Prior Probabilities?

- This is the hardest part, but getting easier in the information age
- **Screening tests** - population prevalence of disease
- **Common Clinical Scenarios** - often can be determined from available decision aides
- **More Complex Scenarios** - might have to guess based on your differential diagnosis
  - Rank the order of likelihood
  - Total = 100%

Examples in your Practice where you can Estimate Probability?

- Screening Tests
- Common Clinical Scenarios
- More Complex Scenarios

Test Results: Sensitivity & Specificity
• You have estimated the probability of disease
• High enough that you cannot reassure the patient
• Low enough that you do not want to just treat empirically
• *A diagnostic test is ordered*...

Tests Should Distinguish Persons With and Without Disease

Population w/o Disease

Population with Disease

Normal Range

Low

Test Levels

High

This would be a bad test... (e.g. CRP)

Population w/o Disease

Population with Disease

Low

Test Levels

High

Let’s focus on the population with disease...

Questions:

- What do we call Group A?
- What do we call Group B?

Of the population with disease, the % that are in the A group is: Sensitivity

Cut point for abnormal test

Normal

Abnormal

Low

Test Levels

High
Now let’s look at the population without disease...

Questions:
- What do we call Group C?
- What do we call Group D?

Of the population without disease, the % that are in the C group is: Specificity

Abnormal

Normal

Low

Test Levels

High

Is your patient more likely to have the disease with this result?

Normal

Abnormal

Low

Test Levels

High

What happens to sensitivity and specificity when we move Cutpoint Lower?

Normal

Abnormal

Low

Test Levels

High

1st KEY POINT: Sensitivity and specificity are about populations, not your patient

Questions:
- Can you tell if your patient has the disease or not?
- If you knew which population the patient was in, then you would not need the test

2nd KEY POINT: Single cut points for tests waste a lot of information

Now your patient gets a test result...
What happens to sensitivity and specificity when we move Cutpoint Higher?

3rd KEY POINT: Moving the cutpoint for an abnormal test trades off sensitivity and specificity.

Prior example makes the disease/no disease populations look about the same size.
That would imply a disease prevalence/prior probability of 50%.
That might be true in a symptomatic patient (e.g. SOB in E.D.).
What about in screening situations, where prevalence is low?

Now, the false positives are a lot more likely than true positives.

4th KEY POINT: The test result has to be interpreted in the context of the probability of disease.

Likelihood Ratios

- Likelihood ratios measure how influential the test result is on the probability of disease in a positive or negative direction.
- LR > 1 = +
- LR < 1 = -
- LR ~ 1 = useless
- Roughly speaking: prior probability x LR = post test probability.
- Actual calculation: pretest odds x LR = post test odds.
So What is a Likelihood Ratio?

- "WOWO" = with over without
  \[
  LR = \frac{\text{Likelihood of result in pop. with disease}}{\text{Likelihood of result in pop. without disease}}
  \]

- **EXAMPLE:** Among patients with dyspnea
  - 60% of those with pneumonia have a productive cough
  - 10% of those without pneumonia have a productive cough

- LR (productive cough) for pneumonia: 60% + 10% = 6

You Can Calculate the Likelihood Ratio for a Dichotomous Test

\[
LR(+) = \frac{\text{Sensitivity}}{1 \text{ minus specificity}}
\]

\[
LR(-) = \frac{1 \text{ minus sensitivity}}{\text{Specificity}}
\]

The LR ruler is a shortcut we can use to get from pre-test probability to post-test probability.

<table>
<thead>
<tr>
<th>Likelihood ratio</th>
<th>Effect of test result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very small (0.1)</td>
<td>Greatly decreases P(disease)</td>
</tr>
<tr>
<td>Less than 1 (0.5)</td>
<td>Decreases P(disease)</td>
</tr>
<tr>
<td>One</td>
<td>No effect on P(disease)</td>
</tr>
<tr>
<td>More than 1 (2)</td>
<td>Increases P(disease)</td>
</tr>
<tr>
<td>Very big (10)</td>
<td>Greatly increases P(disease)</td>
</tr>
</tbody>
</table>
Line-up LR = 1 at Pre-test Probability

Prior probability = 20%

Find LR for Your Test Result

LR (result) = 20

Determine Post-Test Probability

Post-test probability is about 80-85%

Very Good Tests

- All possible results have high or low likelihood ratios

Example:
- HLA B27 for ankylosing spondylitis
  - LR (+) = 15
  - LR (-) = 0.11

- Weak tests- commonly have results that are not definitive (LR close to 1)
- Likelihood ratios can be found in the medical literature
A Chest Pain Example

- 62-year-old man; HTN; chol. = 240 mg/dl
- Chest pain after walking up 2 flights of stairs; radiates to left arm
  - Relieved by resting 5 minutes
  - Gradually progressing
- You decide to order an Exercise Treadmill Test (ETT).

Before the ETT let’s estimate his likelihood of CAD

<table>
<thead>
<tr>
<th>10</th>
<th>25</th>
<th>50</th>
<th>75</th>
<th>90</th>
</tr>
</thead>
</table>

Exercise Treadmill Test

- How do we find the LR+ and LR- for an ETT? (~15 min. search on PubMed)
- Finding the LR:
  *Gianross R. Exercise-induced ST depression in the diagnosis of coronary artery disease. Circulation, 1989*
- Meta-analysis; 24,000 patients, 110 studies
  - LR+ = 3.0
  - LR- = 0.42

Pretest probability of coronary heart disease in patients with chest pain according to age, gender, and symptoms

<table>
<thead>
<tr>
<th>Age</th>
<th>Atypical angina</th>
<th>Typical angina</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Woman</td>
</tr>
<tr>
<td>30-39</td>
<td>34</td>
<td>12</td>
</tr>
<tr>
<td>40-49</td>
<td>51</td>
<td>22</td>
</tr>
<tr>
<td>50-59</td>
<td>65</td>
<td>31</td>
</tr>
<tr>
<td>60-69</td>
<td>72</td>
<td>51</td>
</tr>
</tbody>
</table>

Diamond GA et al., N Engl J Med 1979
Weiner DA et al., N Engl J Med 1979

Septic Arthritis

Bacterial infection in a joint.
Clinical Scenario
Does this Adult Patient Have Septic Arthritis?

A 48-year-old woman with a history of rheumatoid arthritis who has been treated with long-term, low-dose prednisone presents to the emergency department with a 2-day history of a red, swollen right knee that is painful to touch. She reports no prior knee swelling and no recent trauma or knee surgery, illegal drug use, rash, uveitis, or risky sexual behavior. On examination, she is afebrile and has a right knee effusion. Her peripheral white blood cell (WBC) count is 11,000/μL and her erythrocyte sedimentation rate (ESR) is 55 mm/h. An arthrocentesis is performed, and the initial Gram stain is negative.

You have the synovial white blood cell (WBC) count.


Clinical Scenario
Does this Adult Patient Have Septic Arthritis?

According to the review, the pre-test probability of septic arthritis is 38%.

How do you use the synovial WBC result to determine the likelihood of septic arthritis?


Clinical Scenario
Does this Adult Patient Have Septic Arthritis?

Synovial WBC >25K is Considered Positive

<table>
<thead>
<tr>
<th>Synovial WBC Count</th>
<th>Septic Arthritis</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;25,000</td>
<td>77% Sensitivity</td>
</tr>
<tr>
<td>≤ 25,000</td>
<td>23% Specificity</td>
</tr>
</tbody>
</table>

*Note that these could have come from a study where the patients with septic arthritis (D+ patients) were sampled separately from those without (D- patients).

We Can Calculate the Likelihood Ratios

\[ LR(+) = \frac{\text{Sensitivity}}{1 \text{ minus specificity}} \]
\[ LR(-) = \frac{1 \text{ minus sensitivity}}{\text{Specificity}} \]

Sensitivity = 77%
Specificity = 73%

\[ LR(+) = 0.77/(1 - 0.73) = 2.9 \]
\[ LR(-) = (1 - 0.77)/0.73 = 0.32 \]

“+” = > 25,000/uL
“-” = ≤ 25,000/uL

Clinical Scenario
Synovial WBC = 48,000/mL

Pre-test prob: 0.38
LR(+) = 2.9
Post-Test prob = 65%

Use your Slide Rule

What if the Synovial WBC = 128,000/mL

Pre-test prob: 0.38
LR = ?

Should it still be 2.9?

This is an example of where a single cut point throws away information

Synovial Fluid WBC Count
Re-calculate Post-Test Probability

- Pre-test probability: 38%
  
  a) WBC > 25,000, LR= 2.9  
  Expectancy  =65%

  b) WBC > 128,000, LR= 29  
  Probability = ~95% 

  c) WBC = 48,000, LR= 0.8  
  Probability = 33%

Likelihood Ratio

<table>
<thead>
<tr>
<th>WBC (/uL)</th>
<th>% of D+</th>
<th>% of D-</th>
<th>Interval</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;100,000</td>
<td>29%</td>
<td>1%</td>
<td>29.0</td>
<td></td>
</tr>
<tr>
<td>&gt;50,000-100,000</td>
<td>33%</td>
<td>7%</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>&gt;25,000-50,000</td>
<td>15%</td>
<td>19%</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>0 - 25,000</td>
<td>23%</td>
<td>73%</td>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>

An Advantage of LR’s

- They can be used in sequence
  
  Pretest prob (A) x LR¹ = post test prob. (B) 
  
  Prob (B) x LR² = post test prob. (C) 

  OR

- The LRs can all be multiplied together (key shortcut!)
  
  Ex. 4.0 x 0.4 x 5.0 = 8.0 
  (+Test 1) (-Test 2) (+Test 3)
Probability of PE

- 35-year-old man with dyspnea
- Estimated prior probability of PE = 5%
  - Pleuritic pain LR = 3
  - History of DVT LR = 50
  - Patchy infiltrate LR = 0.5
  - pO2 = 92 mm Hg LR = 0.1
- Summary of LR = 3 x 50 x 0.5 x 0.1
  \[\frac{3 \times 25 \times 0.1}{75 \times 0.1} = 7.5\]

Post-Test Probability of PE

\[3 \times 50 \times 0.5 \times 0.1 = 7.5\]

Post-test probability is about 30%

Recommended Reading to Learn More

Thomas B. Newman and Michael A. Kohn

Thank You

Questions?
Example Problems

- Breast cancer
- BNP

Breast Cancer Scenario #1

- A 45 yo woman is contemplating a 1st mammogram. According to National Surveillance Data (SEER) her risk is 0.003 (0.3%). What will her likelihood of breast cancer be if her mammogram report says “additional evaluation needed”? (LR = 7)

- How likely is it that the subsequent needle biopsy will be benign (no cancer)?

Breast Cancer Scenario #2

- A 45 yo woman is contemplating a 1st mammogram. According to National Surveillance Data (SEER) her risk is 0.003 (0.3%). What will her likelihood of breast cancer be if her mammogram report says “suspicious for malignancy”? (LR = 124)

Breast Cancer Scenario #3

- A 65 yo woman is contemplating a 1st mammogram. According to National Surveillance Data (SEER) her risk is 0.013 (1.3%). What will her likelihood of breast cancer be if her mammogram report says “suspicious for malignancy”? (LR = 124)
Breast Cancer
Scenario #4

• A 65 yo woman is contemplating a subsequent mammogram. According to National Surveillance Data (SEER) her risk is 0.003 (0.3%). What will her likelihood of breast cancer be if her mammogram report says “additional evaluation needed”? (LR = 27)

BNP Example

• B-type natriuretic peptide is used to determine whether a patient has heart failure. It’s used commonly in the Emergency Department in patients with dyspnea. The test will be positive (>100) in approximately 90% of patients with underlying HF, and also in approximately 15% of persons without heart failure.

BNP Questions

• What is the sensitivity of the test?
• What is the specificity?
• What is the LR+? LR-?

BNP Questions (continued...)

• How likely is it that your patient has heart failure if:
  ▫ The patient is short of breath in the ER (pre-test probability = 50%)
  ▫ An outpatient known to have an enlarged heart (pre-test probability= 10%)
  ▫ An average middle-aged adult in the community (pre-test probability= 3%)

Surveillance Data (SEER) her risk is
BNP Questions (continued...)

- How would sensitivity and specificity differ if the cutoff for abnormal was 300 instead of 100?