Comparison of Stratum-Specific Likelihood Ratios and 2x2 Approaches

Henry Glick

Epi 550

February 3, 2012

Dismissal of Sensitivity and Specificity

"Readers who have followed the discussion [about likelihood ratios] to this point will understand the essentials of interpretation of diagnostic tests and can stop here. They should consider the next section, which deals with sensitivity and specificity, optional. "We include it largely because clinicians will encounter studies that present their results in terms of sensitivity and specificity and may wish to understand this alternative framework for summarizing properties of diagnostic tests."


Relative Merits of Optimal 2x2 Tables and SSLR for One (Test) and Done Decision Making

• In One and Done decision making, use of the optimal 2x2 table and SSLR yield identical treatment decisions, because in the optimal 2x2 table:
  – All SSLR that yield post-test probabilities above the treatment threshold will be classified as positive tests;
  – All SSLR that yield post-test probabilities below the treatment threshold will be classified as negative tests

NOTE: ALL SSLR, NOT ALL 2x2 TABLES
Relative Merits, Continuous Updating Decision Making

- In continuous updating decision making, use of SSLR will generally yield more discriminating and better calibrated post-test probabilities than use of the optimal 2x2 table
- In addition:
  - No stratum specific result that should yield a post-test probability greater than the pre-test probability will ever yield one that is less than the pre-test probability
  - Very positive and very negative results won’t sometimes have diluted LR associated with them; etc.

Demonstration That for One and Done Decision Making, 2x2 and SSLR Approaches Yield The Same Treatment Decision

- WBC SSLR for Bacteremia

<table>
<thead>
<tr>
<th>Cut-off</th>
<th>SSLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥25</td>
<td>7.6237</td>
</tr>
<tr>
<td>&gt;20, &lt;25</td>
<td>3.0733</td>
</tr>
<tr>
<td>≥15, &lt;20</td>
<td>1.7928</td>
</tr>
<tr>
<td>≥10, &lt;15</td>
<td>0.7920</td>
</tr>
<tr>
<td>≥0, &lt;10</td>
<td>0.1791</td>
</tr>
</tbody>
</table>

- Assume that $C_{fp} = C_{fn}$ → $p^* = 0.50$

Pre-test Probabilities Below the Treatment Threshold

- Post-test probabilities given 3 pre-test probabilities below the treatment threshold and the 5 WBC SSLR

<table>
<thead>
<tr>
<th>Pre-test Probabilities Below the Treatment Threshold</th>
<th>Positive Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.6237</td>
<td>Positive Test</td>
</tr>
<tr>
<td>3.0733</td>
<td>Positive Test</td>
</tr>
<tr>
<td>1.7928</td>
<td>Positive Test</td>
</tr>
<tr>
<td>0.7920</td>
<td>Positive Test</td>
</tr>
<tr>
<td>0.1791</td>
<td>Positive Test</td>
</tr>
</tbody>
</table>
### Agreement Between Post-Test/P* and OOS/Tangency

<table>
<thead>
<tr>
<th>Pre-Test</th>
<th>Strata</th>
<th>OOS</th>
<th>Tangency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>≥ 15</td>
<td>0.6 / 0.4 = 1.500 (0.792; 1.793)</td>
<td>≥ 15</td>
</tr>
<tr>
<td>0.358</td>
<td>≥15 or ≥ 20</td>
<td>0.642 / 0.358 = 1.793 (1.793)</td>
<td>≥15 or ≥ 20</td>
</tr>
<tr>
<td>0.30</td>
<td>≥ 20</td>
<td>0.7 / 0.3 = 2.333 (1.793; 3.073)</td>
<td>≥ 20</td>
</tr>
</tbody>
</table>

### Pre-test Probabilities Below the Treatment Threshold

- Post-test probabilities given 3 pre-test probabilities above the treatment threshold and the 5 WBC SSLR

### Agreement Between Post-Test/P* and OOS/Tangency (2)

<table>
<thead>
<tr>
<th>Pre-Test</th>
<th>Strata</th>
<th>OOS</th>
<th>Tangency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.60</td>
<td>≥ 10</td>
<td>0.4 / 0.6 = 0.667 (0.179; 0.792)</td>
<td>≥ 10</td>
</tr>
<tr>
<td>0.558</td>
<td>≥10 or ≥ 15</td>
<td>0.442 / 0.558 = 0.792 (0.792)</td>
<td>≥10 or ≥ 15</td>
</tr>
<tr>
<td>0.51</td>
<td>≥ 15</td>
<td>0.49 / 0.51 = 0.961 (0.792; 1.793)</td>
<td>≥ 15</td>
</tr>
</tbody>
</table>

*CLp = Ch; (CLp / (CLp + Ch)) = 0.5 = Threshold
*Likelihood ratios defined for WBC for bacteremia
Classifying WBC Counts Between 20 and 25 (SSLR = 3.073) as Positive or Negative Tests

Agreement Between Pre-Test/Post-Test and OOS/Positive/Negative

More than a Demonstration That for One and Done Decision Making, 2x2 and SSLR Approaches Yield The Same Treatment Decision

- Is it possible for 1) a stratum to be classified as a negative test if 2) the post-test probability resulting from its stratum specific likelihood ratio is above the treatment threshold?
1) “Stratum is Classified as a Negative Test”

- Suppose the OOS = 3
- Which of the following strata represent negative results and which positive?

<table>
<thead>
<tr>
<th>Stratum</th>
<th>SSLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>5.0</td>
</tr>
<tr>
<td>Medium</td>
<td>2.5</td>
</tr>
<tr>
<td>Low</td>
<td>0.5</td>
</tr>
</tbody>
</table>

- What does that tell us about the relationship between the OOS and SSLR of strata that are classified as a negative test?

2) Post-test Probability Above the Treatment Threshold

- What do we know about the relationship between the post-test probability from a positive test and the treatment threshold \( P^* \)?

\[
\text{Post-Test} = \frac{\text{SSLR} \cdot p}{(\text{SSLR} \cdot p) + (1-p)} > P^* = \frac{C_p}{C_p + C_n}
\]
Rearranging the Equation, Post-Test > p^*

\[
\frac{\text{SSLR}_p}{(\text{SSLR}_p, p) + (1-p)} > \frac{C_p}{C_p + C_n}
\]

1. Multiply through by the denominators:
   \((\text{SSLR}_p, p C_p) + (\text{SSLR}_p, p C_n) > (\text{SSLR}_p, p C_p) + ((1-p) C_n)\)
2. Cancel \((\text{SSLR}_p, p C_p)\)
   \((\text{SSLR}_p, p C_n) > ((1-p) C_n)\)
3. Divide through by \(p C_n\)
   \(\text{SSLR}_p > \frac{(1-p) C_n}{p C_n}\)

Contradiction

- Classification of a stratum as a negative test implies:
  \(\text{SSLR}_p < \frac{(1-p) C_n}{p C_n}\)
- A post-test probability above the treatment threshold implies:
  \(\text{SSLR}_p > \frac{(1-p) C_n}{p C_n}\)
- Thus, it is impossible for a stratum 1) to be classified as a negative test 2) if the post-test probability resulting from stratum specific likelihood ratio is above the treatment threshold

2) More than a Demonstration That for One and Done Decision Making, 2x2 and SSLR Approaches Yield The Same Treatment Decision

- We can develop an analogous set of equations to show that if a stratum-specific result is classified as a positive test, it cannot yield a post-test probability below the treatment threshold
Summary, "One and Done" Decision Making

- In the optimal 2x2 table, stratum-specific results whose post-test probabilities of disease are above the treatment threshold are classified as positive tests.
- Stratum-specific results whose post-test probabilities of disease are below the treatment threshold are classified as negative tests.
- Thus, in One and Done decision making, in which our treatment decision is based on whether our post-test probability is >p*/<p*, use of optimal 2x2 tables and SSLR yield identical treatment decisions.

Continuous Updating of Probabilities Decision Making

- Use of the optimal 2x2 table generally will yield different post-test probabilities than use of SSLR.
  - The post-test probabilities from the SSLR will generally be better calibrated and more discriminating than those from the optimal 2x2 table.
- Example: IL-6 for bacteremia

<table>
<thead>
<tr>
<th>IL-6 Level</th>
<th>W / Bact</th>
<th>W/O Bact</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥10^3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>≥10^2</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>&lt;10^2</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>46</td>
</tr>
</tbody>
</table>

3 IL-6 Categories, SSLR Approach

<table>
<thead>
<tr>
<th>IL-6 Level</th>
<th>W / Bact</th>
<th>W/O Bact</th>
<th>SSLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥10^3</td>
<td>4</td>
<td>2</td>
<td>4.182</td>
</tr>
<tr>
<td>≥10^2 - &lt;10^3</td>
<td>13</td>
<td>9</td>
<td>3.020</td>
</tr>
<tr>
<td>&lt;10^2</td>
<td>5</td>
<td>35</td>
<td>0.299</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>46</td>
<td>--</td>
</tr>
</tbody>
</table>
2x2 and SSLR Approaches and Post-test Probabilities, IL-6 Example #1

• Because of the combining of strata in the 2x2 approach, the 2x2 and SSLR approaches yield different post-test probabilities of disease

• Suppose the optimal 2x2 table combines $10^2$-$10^3$ and $>10^3$ strata into a positive test (e.g., an OOS of 2.5)
  – SSLR, $>10^3$: 4.182
  – LR+, $>10^2$: $(17 \times 46) / (11 \times 22) = 3.231$
  – SSLR, $10^2$-$10^3$: 3.020

• The post-test probability from the 2x2 approach shifts in the correct direction for both test results
  – However, for $10^2$-$10^3$, the 2x2 approach leads to too great an increase in post-test probability
  – For $>10^3$, it leads to too little an increase
2x2 and SSLR Approaches and Post-test Probabilities, II-6 Example #2

- Suppose the optimal 2x2 table combines $10^2$-$10^3$ and $<10^2$ strata into a negative test (e.g., an OOS of 3.5)
  - SSLR, $10^2$-$10^3$: 3.020
  - LR-, $<10^2$: $(18 \times 46) / (44 \times 22) = 0.855$
  - SSLR, $<10^2$: 0.299
- For test results $<10^2$, the post-test probability from the 2x2 approach shifts in the correct direction (LR-=0.855), although not nearly enough (SSLR=0.299)
- For test results of $10^2$-$10^3$, however, the 2x2 approach shifts the post-test probability in the wrong direction (use of an LR of 0.855 rather than 3.020)

Any Role Left for 2x2 approach?

- Can use SSLR to calculate post-test probabilities that are at least as “good”, if not better, than those derived from the optimal 2x2 table
- Can use either approach to identify “positive” tests / “positive” strata
- Does that mean that Jaeschke, et al. were correct that once you understand SSLR, there is no more role for the selection of the optimal 2x2 table?

Continuous Updating and Multiple Tests

- The purported benefits of the use of SSLR vs the 2x2 approach arise primarily in the context of using multiple tests in the “continuous” updating paradigm
  - Benefits accrue if test results are “independent”
- Multiple tests can be performed:
  - In parallel (multiple tests performed at the same time)
  - In sequence
- Choice affects cost
  - If 1) test results are “independent”, 2) we use SSLR to update probabilities and 3) we perform the same tests either in parallel or sequence, timing of tests will not affect final post-test probability of disease
EKG and Stress Testing for CAD *

<table>
<thead>
<tr>
<th>EKG</th>
<th>Exercise Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D+</td>
</tr>
<tr>
<td>EKG+</td>
<td>256</td>
</tr>
<tr>
<td>EKG-</td>
<td>767</td>
</tr>
<tr>
<td></td>
<td>1023</td>
</tr>
</tbody>
</table>

LR+    = 1.455
LR-    = 0.905

Post-Test Probabilities, Single Tests

- Assume a pre-test probability of disease of 0.2
  - EKG+
    \[
    \frac{0.2 \times 1.455}{0.2 \times 1.455 + 0.8} = 0.267
    \]
  - EKG-
    \[
    \frac{0.2 \times 0.905}{0.2 \times 0.905 + 0.8} = 0.184
    \]
  - Stress+
    \[
    \frac{0.2 \times 3.062}{0.2 \times 3.062 + 0.8} = 0.436
    \]
  - Stress-
    \[
    \frac{0.2 \times 0.275}{0.2 \times 0.275 + 0.8} = 0.064
    \]

Post-Test Probabilities, Multiple Tests

- Multiple ways to calculate post-test probabilities when using more than 1 test
- Pretest → (Test 1) post-test 1 → (Test 2) post-test 2
  - Can use either sensitivity/specificity or LR to estimate post-test probabilities
- Reduce calculations if use multiple LR in one equation
  \[
  \text{pre} \times \text{LR1} \times \text{LR2} \div \left( \text{pre} \times \text{LR1} \times \text{LR2} + (1-\text{pre}) \right)
  \]
- e.g., pre = 0.2; test results: Stress+ / EKG+
  \[
  \frac{0.2 \times 3.062 \times 1.455}{(0.2 \times 3.062 \times 1.455) + 0.8} = 0.527
  \]
- Also applies to OR
  \[
  \frac{\text{punexp} \times \text{OR1} \times \text{OR2}}{(\text{punexp} \times \text{OR1} \times \text{OR2} + (1-\text{punexp}))}
  \]

Post-Test Probabilities, 2 Tests

- Assume a pre-test probability of disease of 0.2
  
  Str+/EKG+ \((0.2*3.062*1.455) / ((0.2*3.062*1.455)+0.8) = 0.527\)
  
  Str+ only \(0.436\)
  
  Str+/EKG- \((0.2*3.062*0.905) / ((0.2*3.062*0.905)+0.8) = 0.409\)
  
  Str-/EKG+ \((0.2*0.275*1.455) / ((0.2*0.275*1.455)+0.8) = 0.091\)
  
  Str- only \(0.064\)
  
  Str-/EKG- \((0.2*0.275*0.905) / ((0.2*0.275*0.905)+0.8) = 0.059\)
  
- For these probabilities to be correct, stress and EKG tests results must be “independent”

### 2 Test Independent SSLR

<table>
<thead>
<tr>
<th>IL-6 Level</th>
<th>Test 1 LR</th>
<th>Test 2 LR</th>
<th>SSLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST+/EKG+</td>
<td>3.062</td>
<td>X</td>
<td>1.455</td>
</tr>
<tr>
<td>ST+/EKG-</td>
<td>3.062</td>
<td>X</td>
<td>0.905</td>
</tr>
<tr>
<td>ST-/EKG+</td>
<td>0.275</td>
<td>X</td>
<td>1.455</td>
</tr>
<tr>
<td>ST-/EKG-</td>
<td>0.275</td>
<td>X</td>
<td>0.905</td>
</tr>
</tbody>
</table>

### What is Test Independence?

- 2 tests are independent if a test’s characteristics remain constant among the other test’s positive and negative findings
  - Defined within disease categories (i.e. has to hold separately for diseased and nondiseased individuals)
  - Needn’t apply to overall sample
- e.g., If test 2 has a specificity of 87.5%:
  - Among those without disease, 87.5% who have a negative test 1 should have a negative test 2; 12.5% should have a positive test 2
  - Among those without disease, 87.5% who have a positive test 1 should have a negative test 2; 12.5% should have a positive test 2
EKG and Stress Testing for CAD Independence?

<table>
<thead>
<tr>
<th>Counts</th>
<th>CAD+</th>
<th>CAD-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stress-</td>
<td>Stress+</td>
</tr>
<tr>
<td>EKG+</td>
<td>32</td>
<td>224</td>
</tr>
<tr>
<td>EKG-</td>
<td>176</td>
<td>591</td>
</tr>
</tbody>
</table>

Row Percentages

Column Percentages

Exact p =0.000

What Independent Tests Might Look Like

<table>
<thead>
<tr>
<th>Counts</th>
<th>CAD+</th>
<th>CAD-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stress-</td>
<td>Stress+</td>
</tr>
<tr>
<td>EKG+</td>
<td>44</td>
<td>212</td>
</tr>
<tr>
<td>EKG-</td>
<td>130</td>
<td>637</td>
</tr>
</tbody>
</table>

Row Percentages

Column Percentages

Exact p=1.0

Agreement and Independence

- Is strong agreement between 2 tests -- i.e., when one is positive the other is almost always positive and when one is negative the other is almost always negative -- suggestive of independence or lack of independence?
- Independence implies informative lack of agreement
  - The strong impact of a false negative result from 1 test will be mitigated by a true positive from the other
  - The strong impact of a false positive result from 1 test will be mitigated by a true negative from the other
- Strong agreement implies the second test generally won't provide mitigating evidence against the mistakes of the first
Test Characteristics Given Dependence

<table>
<thead>
<tr>
<th>IL-6 Level</th>
<th>CAD+</th>
<th>CAD-</th>
<th>SSLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST+/EK+</td>
<td>224</td>
<td>35</td>
<td>2.765 (4.455)</td>
</tr>
<tr>
<td>ST+/EK-</td>
<td>591</td>
<td>80</td>
<td>3.192 (2.771)</td>
</tr>
<tr>
<td>ST-/EK+</td>
<td>32</td>
<td>41</td>
<td>0.337 (0.400)</td>
</tr>
<tr>
<td>ST-/EK-</td>
<td>176</td>
<td>286</td>
<td>0.266 (0.249)</td>
</tr>
<tr>
<td>Total</td>
<td>1023</td>
<td>442</td>
<td>--</td>
</tr>
</tbody>
</table>

Post-Test Probabilities Given Dependence

- Assume a pre-test probability of disease of 0.2

  \[
  \text{ST+/EK+} = \frac{(0.2\times2.765)}{(0.2\times2.765)+0.8} = 0.409 (0.527) \\
  \text{ST+/EK-} = \frac{(0.2\times3.192)}{(0.2\times3.192)+0.8} = 0.444 (0.409) \\
  \text{ST-/EK+} = \frac{(0.2\times0.337)}{(0.2\times0.337)+0.8} = 0.078 (0.091) \\
  \text{ST-/EK-} = \frac{(0.2\times0.266)}{(0.2\times0.266)+0.8} = 0.062 (0.059)
  \]

- Correlations between the 2 tests reduce the gains we’d expect to see if they were actually independent.
- Given dependence, the ST+/EK- post-test probability of 0.444 is higher than that for ST+/EK+.
  - Opposite what we saw under independence assumption

Continuous Updating and Multiple Tests (2)

- If the independence assumption holds, chaining multiple SSLR will yield more discriminating and better calibrated predictions than will use of 2 optimal 2x2 tables.
- For most "paired" tests, not clear how much evidence there is to support independence assumption.
- Lack of independence undermines chaining both the 2x2 and SSLR approaches to updating probabilities.
Continuous Updating and Test Dependence

- Multiple responses one available to address test dependence (e.g., reporting separate SSLR for the single test plus dependent SSLR)
- One response is to treat 2 dependent tests as a single test (in the case of 2 2x2 tables, a single test with 4 outcomes: +/+, +/-, -/+ , --)
- If 2 tests are combined into a single test, what looked like a continuous updating scenario reduces to a one-and-done scenario
- In that case, the selection of the optimal 2x2 table will yield the same results as the use of the dependent SSLR

Are These Tests for AD Independent?

Independence of Tau and Aβ1-42?

<table>
<thead>
<tr>
<th></th>
<th>Alzheimer’s</th>
<th>No Alzheimer’s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aβ+</td>
<td>Aβ-</td>
</tr>
<tr>
<td>Tau+</td>
<td>37 (.661)</td>
<td>2 (.036)</td>
</tr>
<tr>
<td>Tau-</td>
<td>17 (.304)</td>
<td>0 (.000)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Alzheimer’s</th>
<th>No Alzheimer’s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aβ+</td>
<td>Aβ-</td>
</tr>
<tr>
<td></td>
<td>3 (.058)</td>
<td>1 (.019)</td>
</tr>
<tr>
<td></td>
<td>9 (.173)</td>
<td>39 (.75)</td>
</tr>
<tr>
<td></td>
<td>p = 1.0</td>
<td>p = 0.03</td>
</tr>
</tbody>
</table>

- Primary source of failure due to patients without AD
Dependent Tests

- IL-6 and ANC for bacteremia
- EKG and stress tests for CAD
- Tau and Aβ₁₋₄₂ for AD
- CT and Ultrasound for appendicitis
- Mammography and ultrasound for breast cancer
- PfHRP₂ and pLDH-based rapid tests for severe malaria
- BINAXNOW and CARESTART immuno-chromatographic tests and microscopy for malaria
- BD GeneOhm and Cepheid Xpert molecular assays for MRSA
- PCR, DNAP, and culture for chlamydia

Do the Differences Make a Difference?

- If no other test is available or if every SSLR moves you outside the threshold for additional testing
  - The two methods yield the same treatment decision, because they both leave you on the same side of the underlying treatment threshold
  - This conclusion is true even if the likelihood ratios from the 2x2 approach and the SSLR are on opposite sides of 1.0

Do Nothing/Test & Test/Treat Thresholds

- If tests are independent and some stratum-specific results leave you within the testing range and others move you outside it, the 2x2 approach will yield more mistakes
  - SSLR moves you outside testing range, but dilution from 2x2 combination of test results leaves you within the testing range
  - More likely when there are extreme results yielding high SSLR and these SSLR are averaged with other smaller SSLR in the 2x2 approach
Example, Do Nothing/Test Threshold

<table>
<thead>
<tr>
<th>Watchfully wait</th>
<th>Test</th>
<th>Treat empirically</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.25</td>
<td>0.50</td>
</tr>
<tr>
<td>0.75</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Prior = 0.400

IL-6 < 2, SS
IL-6 < 3
LR-
PR = 0.166
0.363

SSLR, IL-6 < 2: 0.299
LR-, IL-6 < 3: 0.855

Test/Treat Thresholds

- Alternatively, SSLR may leave you within the testing range (e.g., 3.020 for IL-6 between 2 and 3), but overstatement of the effect of the test result from the LR+ (3.231 for IL-6 > 2) can move you outside this range

SSLR = 1.0

- Using SSLR, a test can be useful even if some test results yield post-test probabilities equal to our pre-test probabilities (i.e., if one stratum has an SSLR of 1.0)
- In the 2x2 approach, if a positive or negative test result for a particular table yields post-test probabilities equal to our pre-test probabilities:
  1) The other test result does so as well
  2) The ROC curve falls on the 45° line
  3) The test provides no information
### Summary, 2x2 Vs SSLR Approaches

<table>
<thead>
<tr>
<th>2x2 Approach</th>
<th>Stratum-Specific Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Combines proportions of the populations having test results in different strata to develop likelihood ratios for positive and negative tests.</td>
<td>• Does not average among strata (but does average within a stratum).</td>
</tr>
<tr>
<td>➔ For some patients, a test result will yield post-test probabilities that are higher than our prior, while for other patients the same test result will yield post-test probabilities that are lower than one’s prior.</td>
<td>➔ A given test result yields a post-test probability that is either always higher or always lower than our pre-test probability.</td>
</tr>
</tbody>
</table>

---

### Summary, 2x2 Vs SSLR Approaches (2)

<table>
<thead>
<tr>
<th>2x2 Approach</th>
<th>Stratum-Specific Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>• All strata whose results leave us above the treatment threshold will be classified as positive and all strata that leave us below the threshold will be classified as negative.</td>
<td>• Strata always have the same likelihood ratio (Can determine which strata are “positive” and which strata are “negative” by comparison to OOS).</td>
</tr>
<tr>
<td>➔ If the only decision remaining is to treat or withhold treatment, the two approaches yield the same result; if other choices are available, the results can differ.</td>
<td>➔ If the only decision remaining is to treat or withhold treatment, the two approaches yield the same result; if other choices are available, the results can differ.</td>
</tr>
</tbody>
</table>

---

### Summary, 2x2 Vs SSLR Approaches (3)

<table>
<thead>
<tr>
<th>2x2 Approach</th>
<th>Stratum-Specific Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Retains concept of positive test.</td>
<td>• Retains concept of positive test result if we compare the SSLR to the OOS.</td>
</tr>
<tr>
<td>• Cost of mistakes and thresholds built into the definition of a positive test.</td>
<td>• Cost of mistakes and thresholds built into the definition of a “positive” stratum.</td>
</tr>
<tr>
<td>• Can use Bayes theorem or likelihood ratio approach to adjust pre-test probabilities.</td>
<td>• Can use Bayes theorem or likelihood ratio approach to adjust pre-test probabilities.</td>
</tr>
</tbody>
</table>
Summary, 2x2 Vs SSLR Approaches (4)

<table>
<thead>
<tr>
<th>2x2 Approach</th>
<th>Stratum-Specific Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Withhold testing for therapeutic decisions if no stratum-specific result can</td>
<td>(IF One (Test) and DONE) Withhold testing for therapeutic decisions if no stratum-specific result can</td>
</tr>
<tr>
<td>shift the post-test probability and pre-test to opposite sides of the</td>
<td>shift the post-test probability and pre-test to opposite sides of the underlying treatment</td>
</tr>
<tr>
<td>underlying treatment threshold</td>
<td>threshold</td>
</tr>
</tbody>
</table>

Summary, 2x2 Vs SSLR Approaches (5)

<table>
<thead>
<tr>
<th>2x2 Approach</th>
<th>Stratum-Specific Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>So long as there is a mapping function between the optimal operating point</td>
<td>Ideally, the likelihood ratios are based on the proportions of tests among diseased and</td>
</tr>
<tr>
<td>on the maximum likelihood estimate of the ROC curve and the criteria for a</td>
<td>nondiseased among infinitesimally small strata (i.e., on the relationships between</td>
</tr>
<tr>
<td>positive test, the 2x2 approach does not have to prespecify particular cut-</td>
<td>the two density functions)</td>
</tr>
<tr>
<td>points for the diagnostic test results</td>
<td></td>
</tr>
</tbody>
</table>

Take Home Messages

1) We base decisions estimates of the probability of disease and one's treatment thresholds
2) We perform diagnostic tests to change our probabilities / certainty about the appropriate treatment
3) We use a test's sensitivity and specificity or its likelihood ratios to revise our pre-test probability of disease to yield a post-test probability based on whether a test's results are positive or negative or based on its stratum-specific likelihood ratios
4) Different sensitivities and specificities may be appropriate for different patients
Take Home Messages (cont.)

5) Combining data from ROC curves OR SSLR with the OOS represents a good method for identifying the appropriate sensitivity and specificity for a given patient.

6) In One and Done decision making, the optimal sensitivity and specificity yields a decision identical to the one based on stratum-specific likelihood ratios; in continuous updating decision making, use of stratum-specific results yields superior post-test probabilities of disease.