Comparison of Stratum-Specific Likelihood Ratios and 2x2 Approaches

Henry Glick

Epi 550

February 12, 2020

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."

Jaeschke, et al. JAMA. 1994;271:703-707

2 Types of Decision Making

- Previously introduced "one (test) and done," 1 of (at least) 2 types of decision making
  - One and done decision making characterized by need/desire to make treatment decision quickly without a large number of opportunities to collect additional data
- Second type of decision making referred to as "continuous updating"
  - Continuous updating characterized by performing a number of different tests and regularly updating probability of disease based on each test's results
- Today focus on strengths and weakness of 2x2 and SSLR approaches for these 2 decision making paradigms

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

- In One and Done decision making, use of OPTIMAL 2x2 table and SSLR yield identical treatment decisions, because for optimal 2x2 table:
  - All strata with SSLR yielding post-test probabilities above treatment threshold (p*) are classified as positive test results;
  - All strata with SSLR yielding post-test probabilities below treatment threshold are classified as negative test results

NOTE: ALL SSLR YIELDING POST-TEST PROBABILITIES ABOVE P*, NOT ALL 2x2 TABLES

SSLR, Not 2x2 Tables!!!

Cut-off	Stratum	Bact	No Bact	LR+	SSLR
≥25	≥25	6	26	7.624	7.624
≥20	≥20, <25	4	43	4.788	3.073
≥15	≥15, <20	7	129	2.837	1.793
≥10	≥10, <15	7	292	1.618	0.792
≥0 *	≥0, <10 *	2	369	--	0.179
Total	Total	26	859		

- Suppose that p=0.2, $\Delta O_{D+} = \Delta O_{D-}$: OOS = 4 and p* = 0.5
- ≥25 included as a positive test because 1) LR+ of T+ ≥25 and SSLR of stratum ≥25 are both greater than 4 (OOS)
- And 2) both yield a post-test probability > 0.5 (p*)

$(0.2 * 7.624)/((0.2 * 7.624) + 0.8) = 0.66 > 0.5$

SSLR, Not 2x2 Tables!!!

Cut-off	Stratum	Bact	No Bact	LR+	SSLR
≥25	≥25	6	26	7.624	7.624
≥20	≥20, <25	4	43	4.788	3.073
≥15	≥15, <20	7	129	2.837	1.793
≥10	≥10, <15	7	292	1.618	0.792
≥0 *	≥0, <10 *	2	369	--	0.179
Total	Total	26	859		

- When p=0.2, $\Delta O_{D+} = \Delta O_{D-}$, OOS = 4 and p* = 0.5, LR+ of T+ ≥20 also yields post-test probability > 0.5

$(0.2 * 4.788) / (0.2 * 4.788) + 0.8) = 0.54 > 0.5$

  - Does that mean ≥20, <25 stratum should be considered a positive test result?

Should ≥20, <25 Be Considered a Positive Test?

- NO!
- While LR+ (4.788) for T+ ≥20 raises post-test probability above 0.5…

  (0.2*4.788) / ((0.2*4.788) + 0.8) = .5438 > 0.5
- SSLR (3.073) for ≥20, <25 does not

  (0.2*3.073) / ((0.2*3.073) + 0.8) = .4345 < 0.5
- 2x2 table with ≥20 cut-off yields a post-test probability above 0.5 because information it borrows from ≥25 stratum is highly predictive
- NOT because information from ≥20/<25 stratum is highly predictive

Reiteration

Method for identifying optimal 2x2 table ensures classification of strata with SSLR yielding post-test probabilities above treatment threshold as positive tests

Also ensures classification of strata with SSLR yielding post-test probabilities below treatment threshold as negative tests

Demonstration: In One and Done Decision Making, 2x2 and SSLR Approaches Yield Same Decision

- WBC SSLR for Bacteremia

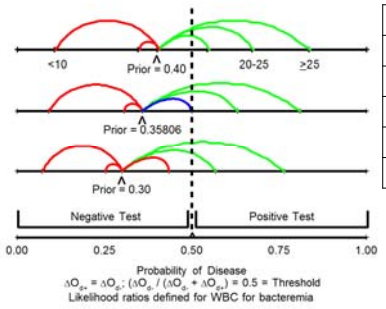
Strata	SSLR
≥25	7.6237
≥20, <25	3.0733
≥15, <20	1.7928
≥10, <15	0.7920
≥0, <10	0.1791

- Assume that $\Delta O_{D+} = \Delta O_{D-} \rightarrow p^* = 0.50$

Pre-test Probabilities Below Treatment Threshold

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

SSLR
7.6237
3.0733
1.7928
0.7920
0.1791

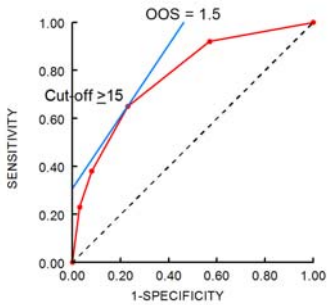


OOS
.6/.4=1.5
1.7928
.7/.3=2.3

Proof: For one and Done Decision Making, 2x2 and SSLR Approaches Yield Same Treatment Decision

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

What mathematical expression is equivalent to "Stratum is classified as a negative test"?



1) "Stratum is Classified as a Negative Test"

- Test results from strata with SSLR less than OOS interpreted as negative

$$SSLR_i < OOS$$

$$OOS = \frac{(1-p)\,\Delta O_{D-}}{p\,\Delta O_{D+}}$$

$$SSLR_i < OOS = \frac{(1-p)\,\Delta O_{D-}}{p\,\Delta O_{D+}}$$

What mathematical expression is equivalent to "Post-test probability is above treatment threshold"?

2) Post-test Probability Above Treatment Threshold

1) $\text{Post-Test} > P^*$

2) $$\text{Post-Test} = \frac{SSLR_i\ p}{(SSLR_i\ p) + (1-p)}$$

3) $$P^* = \frac{\Delta O_{d-}}{\Delta O_{d-} + \Delta O_{d+}}$$

4) $$\frac{SSLR_i\ p}{(SSLR_i\ p) + (1-p)} > \frac{\Delta O_{d-}}{\Delta O_{d-} + \Delta O_{d+}}$$

Rearranging Post-Test > p* Equation

$$\frac{SSLR_i\ p}{(SSLR_i\ p) + (1-p)} > \frac{\Delta O_{D-}}{\Delta O_{D-} + \Delta O_{D+}}$$

1. Multiply through by the denominators:

$$(SSLR_i\ p\ \Delta O_{D-}) + (SSLR_i\ p\ \Delta O_{D+}) > (SSLR_i\ p\ \Delta O_{D-}) + ((1-p)\ \Delta O_{D-})$$

2. Cancel $(SSLR_i\ p\ \Delta O_{D-})$:

$$(SSLR_i\ p\ \Delta O_{D+}) > (1-p)\ \Delta O_{D-}$$

3. Divide through by $p\Delta O_{D+}$

$$SSLR_i > \frac{(1-p)\ \Delta O_{D-}}{p\ \Delta O_{D+}}$$

Contradiction

- Classification of stratum as negative test implies:

$$SSLR_i < \frac{(1-p)\ \Delta O_{D-}}{p\ \Delta O_{D-}}$$

- Post-test probability above treatment threshold implies:

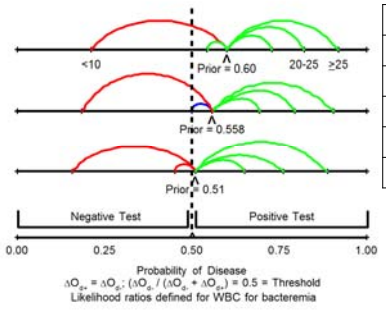
$$SSLR_i > \frac{(1-p)\ \Delta O_{D-}}{p\ \Delta O_{D+}}$$

- Thus, impossible for stratum 1) to be classified as negative test 2) if post-test probability resulting from stratum specific likelihood ratio is above treatment threshold

Pre-test Probabilities Above Treatment Threshold

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

SSLR
7.6237
3.0733
1.7928
0.7920
0.1791



OOS
.4/.6=0.67
0.7920
0.961

Continuation of Proof: For one and Done Decision Making, 2x2 and SSLR Approaches Yield Same Treatment Decisions

- Can develop an analogous set of equations to show that a stratum-specific result classified as positive test, cannot yield post-test probability below treatment threshold
- Starts with:

1) $$SSLR_i > \frac{(1-p)\,\Delta O_{D-}}{p\,\Delta O_{D+}}$$

2) $$\frac{SSLR_i\ p}{(SSLR_i\ p)+(1-p)} < \frac{\Delta O_{D-}}{\Delta O_{D-} + \Delta O_{D+}}$$

Summary, "One and Done" Decision Making

- In optimal 2x2 table, strata that yield post-test probabilities of disease above treatment threshold classified as positive tests
- Strata that yield post-test probabilities below treatment threshold classified as negative tests
- Thus, in One and Done decision making -- in which our treatment decision is based on whether post-test probability is >p*/<p* -- use of optimal 2x2 table and SSLR yield identical treatment decisions

Relative Merits, Continuous Updating Decision Making (1)

1) Use of SSLRs generally yields different post-test probabilities than use of optimal 2x2 table
2) Use of SSLRs generally yields more DISCRIMINATING post-test probabilities than optimal 2x2
   - Why more discriminating?
     - Because SSLR – with no averaging of strata – yield more very high and very low post-test probabilities, whereas 2x2 approach – with more averaging across strata – have more post-test probabilities clustered around pre-test probability

Relative Merits, Continuous Updating Decision Making (2)

2) Unlike the 2x2 approach, in SSLR approach, no stratum-specific result that yields a post-test probability greater than pre-test probability will ever yield one that is less than pre-test probability (and vice versa)

IL-6 Categories for Bacteremia *

IL-6 Level	W / Bact	W/O Bact
$\geq 10^3$	4	2
$\geq 10^2$ - $<10^3$	13	9
$<10^2$	5	35
Total	22	46

* Strait RT, et al. Tumor necrosis factor-a, interleukin-1b, and interleukin-6 levels in febrile, young children with and without occult bacteremia. Pediatrics 1999; 104: 1321–1326

3 IL-6 Categories, 4 IL-6 2x2 Tables

IL-6 Level	W / Bact	W/O Bact	LR+	LR-
All negative	0 22	0 46	--	1.000
$\geq 10^3$	4 18	2 44	4.182	0.855
$\geq 10^2$	17 5	11 35	3.231	0.299
All positive	22 0	46 0	1.000	--

3 IL-6 Categories, 3 SSLR

IL-6 Level	W / Bact	W/O Bact	SSLR
$\geq 10^3$	4	2	4.182
$\geq 10^2$ - $<10^3$	13	9	3.020
$<10^2$	5	35	0.299
Total	22	46	--

Summary of LR Using 2x2 and SSLR Approaches

	Actual Test Result		
2x2 Cut-off	$\geq 10^3$	$\geq 10^2$; $<10^3$	$<10^2$
All negative	1.0	1.0	1.0
$\geq 10^3$	4.182	0.855	0.855
$\geq 10^2$	3.231	3.231	0.299
All positive	1.0	1.0	1.0
SSLR Approach	4.182	3.020	0.299

Suppose Optimal Table Combines 10^2-10^3 and $\geq 10^3$

	Actual Test Result		
2x2 Cut-off	$\geq 10^3$	$\geq 10^2$; $<10^3$	$<10^2$
All negative	1.0	1.0	1.0
$\geq 10^3$	4.182	0.855	0.855
$\geq 10^2$	3.231	3.231	0.299
All positive	1.0	1.0	1.0
SSLR Approach	4.182	3.020	0.299

1) LR+/LR- and SSLR agree about direction of probability shift
2) But 2x2 approach yields too small an increase for $\geq 10^3$ (3.231 vs 4.182) and too large an increase in probability for 10^2-10^3 (3.231 vs 3.020)

Instead Suppose Optimal Table Identifies 10^3 as Test+

	Actual Test Result		
2x2 Cut-off	$\geq 10^3$	$\geq 10^2$; $< 10^3$	$< 10^2$
All negative	1.0	1.0	1.0
$\geq 10^3$	4.182	0.855	0.855
$\geq 10^2$	3.231	3.231	0.299
All positive	1.0	1.0	1.0
SSLR Approach	4.182	3.020	0.299

1) For $<10^2$, 2x2 approach shifts probability in correct direction, but not nearly enough (LR-=0.855; SSLR=0.299)
2) For 10^2-10^3, 2x2 approach shifts post-test probability in wrong direction (LR- of 0.855; SSLR=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 (i.e., more discriminating), than those derived from 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 selection of optimal 2x2 table?

Maybe Not, But…

- Already have seen that SSLR throw away information that aides in the choice between tests
- While many have pointed to SSLR's more effective combination of multiple tests, i.e., continuous updating of probabilities), is continuous updating as unproblematic as the "evidence-based" commentators suggest?

Continuous Updating and Multiple Tests

- Standard model of continuous updating assumes starting with a prior; obtaining test result; updating probability; obtaining another test result; updating;....
  - If want to use a single set of LR for each test whether or not predictions have already been made with other tests, requires that results from multiple tests be "independent"

Timing of Multiple Tests

- Multiple tests can be performed:
  - In parallel (multiple tests performed at same time)
  - In sequence
- All else equal, in parallel/in sequence choice affects cost, but not predicted probability
  - In parallel commits us to performing all tests
  - For "in sequence," if 1) test results are "independent", 2) we use SSLR to update probabilities, same test results will yield same post-test probabilities
    - But may be sufficient information from first few tests so later tests can be avoided

EKG and Stress Testing for CAD Among Men *

	D+	D-
EKG+	256	76
EKG-	767	366
	1023	442

	D+	D-
Stress+	815	115
Stress-	208	327
	1023	442

LR+	1.455	3.062
LR-	0.905	0.275

* Weiner DA, et al. Correlations among history of Angina, ST-segment response, and prevalence of coronary artery disease in the Coronary Artery Surgery Study (CASS). NEJM. 1979;301:230-5.

Continuous Updating: CAD Risk

- Identify pre-test probability (e.g., 20%)
- Order EKG, result is positive, post-test probability:
  - (0.2*1.455) / ((.2*1.455)+0.8) = 0.2667
- Order exercise stress test, result is positive
  - (0.2667*3.062) / ((0.2667*3.062)+0.7333) = 0.5269
- Would also obtain 0.5269 if we ordered stress test first
  - (0.2*3.062) / ((0.2*3.062)+0.8) = 0.4336
- Ordered EKG second
  - (0.4336*1.455) / ((0.4336*1.455)+0.5664) = 0.5269
- Would also obtain 0.5269 if we multiplied both LR
  - (0.2*1.455*3.062) / ((0.2*1.455*3.062)+0.8) = 0.5269
  - Multiplication of LR considered an advantage of SSLR

Why Not Consider 2 Tests as Joint Test?

- 4 potential joint test results (strata)
  - ST+/EK+, ST+/EK-, ST-/EK+, and ST-/EK-
- As we know, 4 strata, 5 2x2 tables
  - All negative
  - ST+/EK+ positive
  - ST+/EK+ and ST+/EK- positive
  - ST+/EK+ and ST+/EK- and ST-/EK+ positive
  - All positive
- If tests independent, can use multiplication to calculate LR+/LR- for the five potential combinations of positive and negative test results

Can Also Calculate SSLR for 4 Combinations of 2 test's results

Effective SSLR, 2 Independent DichotomousTests

ST/EK result	EKG LR		Stress LR	SSLR
ST+/EK+	1.455	X	3.062	4.455
ST+/EK-	0.905	X	3.062	2.771
ST-/EK+	1.455	X	0.275	0.400
ST-/EK-	0.905	X	0.275	0.249

- Results ordered by descending SSLR
- Comparison of OOS to resulting "joint" SSLR indicates which combinations should be considered positive
  - e.g., if OOS= 0.3, ST+/EK+, ST+/EK-, and ST-/EK+ should be included in a "positive" test

Issues With Treating 2 Tests Jointly

- Use of combination of tests as a "joint" test returns us to one and done decision making rather than continuous updating
  - As with ordinary one and done decision making will reach same treatment decisions by use of either optimal combined 2x2 table or test's SSLR
- If want to calculate joint test characteristics using LR / Sens & Spec for each test when used separately, requires test independence
- **But Are EKG and stress tests independent?**

TEST INDEPENDENCE

Independence Similar to Spectrum Bias

- Recall that spectrum bias absent when:
  - Sensitivity homogeneous among D+ patients with and without signs and symptoms
  - Specificity homogeneous among D- patients with and without signs and symptoms
- Spectrum bias present when either is heterogenous

	D+			D-	
	Cent 0/1	Cent 3/4		Cent 0/1	Cent 3/4
RADT+	181	187	RADT+	44	43
RADT-	66	31	RADT-	813	311
	247	218		857	354
Sens	73.3	85.8	Spec	94.9	87.9
p =	0.001			0.000	

Similarities Between Spectrum Bias and Independence

- Evaluation of spectrum bias and test independence have a number of similarities
  - For both assessments, evaluation conducted separately, one 2x2 table among those with disease and one 2x2 table among those without disease
  - Rows represent test results (T+ and T-)
  - Statistically significant differences indicate evidence for presence of spectrum bias AND for lack of independence
    - Spectrum bias / lack of independence present when either sensitivities or specificities differ significantly

Primary Difference

- For spectrum bias, columns represent presence or absence of characteristic for which spectrum bias is being assessed, e.g.:
  - Patients with symptoms vs those without
  - Females vs males
  - Old vs young, etc.
- For test of independence columns represent test results (T+ and T-) for second test

Spectrum Bias Vs Test of Independence, D+

	D+ Cent 0/1	D+ Cent 3/4		D+ Stress+	D+ Stress-
Rapid+	181	187	EKG+	32	224
Rapid-	66	31	EKG-	176	591
	247	218		208	815
Sens	73.3	85.8	Spec	0.154	0.275
	p = 0.001			Exact p=0.000	

- For spectrum bias, is sensitivity same for characteristic such as patients without and with symptoms?
- For independence, is sensitivity of EKG same for patients who are stress+ (stress TP) and stress- (stress FN)?

Spectrum Bias Vs Test of Independence, D-

	D- Cent 0/1	D- Cent 3/4		D- Stress+	D- Stress-
Rapid+	44	43	EKG+	41	35
Rapid-	813	311	EKG-	286	80
	857	354		327	115
Sens	94.9	87.9	Spec	0.875	0.696
	p = 0.000			Exact p=0.000	

- For spectrum bias, is specificity same for characteristic such as patients without and with symptoms
- For independence, is specificity of EKG same for patients who are stress+ (stress FP) and stress- (stress TN)

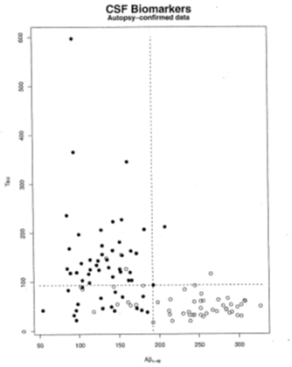
Are EKG and Stress Tests for CAD Independent?

Counts	CAD+ Stress-	CAD+ Stress+		CAD- Stress-	CAD- Stress+
EKG+	32	224	EKG+	41	35
EKG-	176	591	EKG-	286	80
	208	815		327	115
EKG Sens	0.154	0.275	EKG Spec	0.875	0.696
	Exact p =0.000			Exact p=0.000	

What Independent Tests Might Look Like

Counts	CAD+			CAD-	
	Stress-	Stress+		Stress-	Stress+
EKG+	52	204	EKG+	56	20
EKG-	156	611	EKG-	271	95
	208	815		327	115
Sens	0.25	0.25	Spec	0.829	0.826
	p=1.000			p=1.000	

Are These Tests for AD Independent?



Y axis: CSF tau concentration

X axis: CSF amyloid-β 1 to 42 peptide

Solid circles confirmed AD

Open circles: elderly cognitively normal subjects

Independence of Tau and $A\beta_{1-42}$?

	Alzheimer's			No Alzheimer's	
	Aβ+	Aβ-		Aβ+	Aβ-
Tau+	37	2		3	1
Tau-	17	0		9	39
SE	68.5%	100%	SP	75%	97.5%
	p = 1.0 (exact) *			p = 0.03 (exact) †	

- Lack of independence due to very diffèrent proportions of Tau+ and Tau- among Aβ+ (3/12) and Aβ- patients (1/40)

* tabi 37 2 \ 17 0,exact
† tabi 3 1 \ 9 39,exact

Agreement and Independence

- Is strong agreement between 2 tests suggestive of independence or lack of independence?
  - By strong agreement, we mean:
    - i.e., When one test is positive other is almost always positive
    - When one is negative other is almost always negative

Strong Agreement…

- …suggestive of **LACK** of independence
  - Independence holds if, for example, among diseased individuals fraction of test 1's true positives equal among test 2's true positives and false negatives
  - But if tests demonstrate strong agreement:
    - When test 2's results are positive, test 1's results also likely to be positive
      - i.e., too few positives among test 1's false negatives
    - When test 2's results are negative, test 1's results also likely to be negative
      - i.e., too few negatives among test 1's false positives

Informative Lack of Agreement

- Independence implies "informative lack of agreement"
  - Want impact of false negative result from 1 test to be offset by true positive from second
  - Want impact of false positive result from 1 test to be offset by true negative from second
- Strong agreement implies both tests will generally provide same misinformation
  - i.e., both false negative or both false positive

Implication of Dependence for Continuous Updating

- Lack of independence undermines chaining both 2x2 and SSLR approaches for updating probabilities
  - i.e., Can't multiply EKG's LR- times either Stress LR
- One response to test dependence is to treat 2 dependent tests as a single test
- Again, four potential joint test results
  - ST+/EK+, ST+/EK-, ST-/EK+, and ST-/EK-
- Unlike independent tests, don't calculate combined SSLR / LR+/LR- by multiplying each test's SSLR / LR+/LR-
- Instead return to original data and calculate combined SSLR / LR+/LR- from combined results of 2 tests

Observed (Dependent) Vs Calculated (Independent) SSLR, Combined Tests *

ST/EK result	D+	D-	Dep SSLR	Indep SSLR	P-value †
ST+/EK+	224	35	2.765	4.455	0.0004
ST+/EK-	591	80	3.192	2.771	0.02
ST-/EK+	32	41	0.337	0.400	0.33
ST-/EK-	176	286	0.266	0.249	0.04
Total	1023	442	--	--	--

- See data from slide 47

† P-values derived via bootstrap

Continuous Updating and Test Dependence (2)

- As with independent tests, treating 2 dependent tests as a single joint test reduces what looked like continuous updating to one-and-done decision making
  - Selection of optimal joint test 2x2 table yields same treatment decision as SSLR
  - Purported advantage of SSLR may disappear

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 2x2 approach, if a positive (negative) test result for a particular table yields post-test probabilities equal to our pre-test probabilities:
  1) Negative (positive) test result for same table does so as well
  2) ROC curve falls on 45° line
  3) Test provides no information

Summary, 2x2 Vs SSLR Approaches

2x2 Approach	Stratum-Specific Approach
● Combines proportions of populations having test results in different strata to develop likelihood ratios for positive and negative tests	● Does not average among strata (but does average within a stratum)
➔ For some patients, a test result yields post-test probabilities that are higher than pre-test, while for others same test result yields post-test probabilities that are lower than pre-test	➔ A given test result yields post-test probability that is either always higher or always lower than pre-test probability

Summary, 2x2 Vs SSLR Approaches (2)

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

Summary, 2x2 Vs SSLR Approaches (3)

2x2 Approach	Stratum-Specific Approach
● Retains concept of positive test	● Retains concept of positive test result if we compare SSLR to OOS
● Cost of mistakes and thresholds built into definition of a positive test	● Cost of mistakes and thresholds built into definition of a "positive" stratum
● Can use Bayes theorem or likelihood ratio approach to adjust pre-test probabilities	● Can use likelihood ratio approach to adjust pre-test probabilities

Summary, 2x2 Vs SSLR Approaches (4)

2x2 Approach	Stratum-Specific Approach
● Withhold testing for therapeutic decisions if no stratum-specific result can shift post-test probability and pre-test to opposite sides of underlying treatment threshold	● (IF One (Test) and DONE) Withhold testing for therapeutic decisions if no stratum-specific result can shift post-test and pre-test to opposite sides of underlying treatment threshold

Summary, 2x2 Vs SSLR Approaches (5)

2x2 Approach	Stratum-Specific Approach
● Does not have to prespecify particular cut-points for diagnostic test results (i.e., identify cut-off and Se and Sp associated with tangency of OOS and ROC curve)	● Does not have to prespecify particular cut-points for diagnostic test results (i.e., identify cut-off associated with point on OOS that has a slope equal to OOS)

Take Home Messages

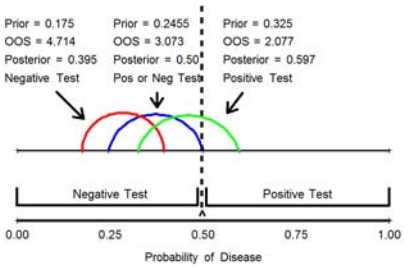
1) We base decisions on estimates of probability of disease and one's treatment thresholds
2) We perform diagnostic tests to change our probabilities / certainty about appropriate treatment
3) We use a test's sensitivity and specificity or its likelihood ratios to revise our pre-test probability of disease 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 OOS represents a good method for identifying appropriate sensitivity and specificity for a given patient
6) In One and Done decision making, optimal sensitivity and specificity yields a decision identical to one based on stratum-specific likelihood ratios; in continuous updating decision making, use of stratum-specific results yields superior post-test probabilities of disease

Extra Slides

Classifying WBC Counts Between 20 and 25 (SSLR = 3.073) as Positive or Negative Tests



Cfp = Cfn; (Cfp / (Cfp + Cfn)) = 0.5 = Threshold
Likelihood ratios defined for WBC for bacteremia

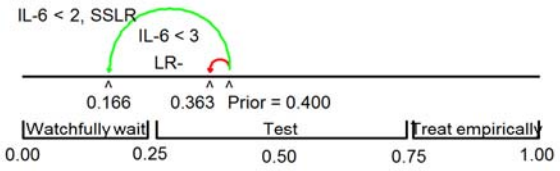
Do Differences Make a Difference?

- If no other test is available or if every SSLR moves you outside threshold for additional testing
  - Two methods yield same treatment decision, because they both leave you on same side of underlying treatment threshold
  - Conclusion true even if likelihood ratios from 2x2 approach and 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 testing range and others move you outside it, 2x2 approach will yield more mistakes
  - SSLR moves you outside testing range, but dilution from 2x2 combination of test results leaves you within testing range
    - More likely when there are extreme results yielding high SSLR and these SSLR are averaged with other smaller SSLR in 2x2 approach

Example, Do Nothing/Test Threshold



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

Test/Treat Thresholds

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

Are EKG and Stress Tests for CAD Independent?

Counts

	CAD+		CAD-	
	Stress-	Stress+	Stress-	Stress+
EKG+	32	224	41	35
EKG-	176	591	286	80
	Sens 0.154	0.275	Spec 0.875	0.696
	Exact p =0.000		Exact p=0.000	

Row Percentages
Column Percentages

12.5 **15.4**	87.5 **27.5**	54.0 **12.5**	46.0 **30.4**
23.0 **84.6**	77.0 **72.5**	78.1 **87.5**	21.9 **69.6**