Principles of Screening

Jeffrey Murawsky, MD
Assistant Professor, Medicine

Perfect Screening Test

- Always correct
- Repeatable
- Safe, painless, quick, inexpensive
- Makes a clinical difference

Reality is Quite a Different Prospect!

Basic Two by Two

Test	Gold Standard: Disease Positive	Gold Standard: Disease Negative
Positive	True Positives	False Positives
Negative	False Negatives	True Negatives

Sensitivity

- Proportion of those with disease defined by gold standard testing who are labeled by the test in question as positive
- True positives/ all subjects with gold standard proven disease

Specificity

- Proportion of those without disease defined by gold standard testing who are labeled as negative by the test in question
- True Negatives/ all subjects disease free by gold standard testing

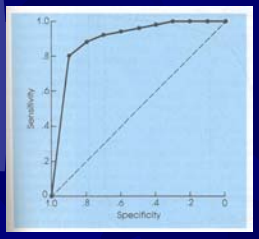
Prevalence

- In the 2x2 table: the number of those with disease by gold standard ie 5/100,000
- Clinically Pretest Probability of the patient is very similar (ie 30% chance of a disease based on risk factors)

Defining a Positive Test

- Tests are usually yield a continuous variable
- An artificial cut off is needed to define the positive or abnormal values from the normal or negative values.
- The Receiver operating characteristic curve demonstrates the trade off

ROC Curve



- As specificity is increased sensitivity is lost.
- The closer to the upper left corner your values yield the better balance in the test characteristics.

Example

ND Disease: " an uncontrollable urge to watch a football team without any hope of winning (especially bowl games)"
We know from research at LUMC that in our medical school this disease occurs in 1 in 10 medical students

Jeff's Test is questionnaire available and has defined sensitivities and specificities by population testing

Given 2000 medical students	Have ND	Don't Have ND
Jeff's Test Positive	True Positives	False Positives
Negative	False Negatives	True Negatives
	200 with ND	1800 without ND

Jeff's Test

Sensitivity: 80%
Specificity: 90%

That means: 80% of 200 or 160 will be True Positives and
90% of 1800 or 1620 will be True Negatives
Thru the magic of Mathematics:

Given 2000 medical students	Have ND	Don't Have ND
Jeff's Test Positive	True Positives 160	False Positives 180
Negative	False Negatives 20	True Negatives 1620
Prevalence of 1 in 10 =	200 with ND	1800 without ND

Jeff's Test

Sensitivity: 80%
Specificity: 90%

Now let's assume that the prevalence of ND disease changes to 1 in 200 in another population.

Given 2,000 persons in Boston	Have ND	Don't Have ND
Jeff's Test Positive	True Positives 4	False Positives 199
Negative	False Negatives 6	True Negatives 1791
	10 with ND	1990 without ND

What's the Most Important Clinical Question?

If the test is positive, what is the chance that the patient really has disease

or

If the test is negative, what is the chance that the patient does not have the disease?

Predictive Value

Positive Predictive Value

- The proportion of patients testing positive who actually have the disease (by gold standard)
- True Positives / All positives

Predictive Value

Negative Predictive Value

- The Proportion of patients testing negative who are truly free of the disease (by gold standard)
- True Negatives / All Negatives

Jeff's Test Sensitivity: 80% Specificity: 90%

Positive Predictive Value: 160 True Positives / 160 + 180 (all testing positive)

Negative Predictive Value: 1620 True Negative/ 1620 + 20 (all testing negative)

Given 2000 medical students	Have ND	Don't Have ND	
Jeff's Test Positive	True Positives 160	False Positives 180	340 Tested Positive
Jeff's Test Negative	False Negatives 20	True Negatives 1620	1640 Tested Negative
Prevalence of 1 in 10 =	200 with ND	1800 without ND	

Jeff's Test Sensitivity: 80% Specificity: 90%

The test is only right about a positive result 47.% of the time

The test is right about a negative result 99% of the time

Given 2000 medical students	Have ND	Don't Have ND	
Jeff's Test Positive	True Positives 160	False Positives 180	340 Tested Positive
Jeff's Test Negative	False Negatives 20	True Negatives 1620	1640 Tested Negative
Prevalence of 1 in 10 =	200 with ND	1800 without ND	

Jeff's Test Sensitivity: 80% Specificity: 90%

PPV: 4/203 =.02 or 2%

NPV: 1791/1797 =.99 or 99%

Given 2,000 persons in Boston	Have ND	Don't Have ND
Jeff's Test Positive	True Positives 4	False Positives 199
Jeff's Test Negative	False Negatives 6	True Negatives 1791
	10 with ND	1990 without ND

A quick short cut:

- As prevalence increases: PPV increases and NPV decreases
- As prevalence decreases: PPV decreases and NPV increases

Given a Test with a sensitivity of 80% and a specificity of 90%:

	Pretest Probability 1%	10%	50%	90%
Positive Predictive Value	7.5%	47.1%	88.9%	98.6%
Negative Predictive Value	99.8%	97.6%	81.8%	33.3%

The point is . . .

- Sensitivity and Specificity are functions of the operating curves of the test
- Predictive Values are related to prevalence or pre test probabilities
- Statistical difference doesn't necessarily relate to clinical relevance
- Clinically Rule In/ Rule Out may not be accomplished by one test

Now let's assume that two tests are now available to screen for a new disease (Barrister's Syndrome) and if detected in the asymptomatic phase can be cured with minimal therapy. The rate of Barrister's Syndrome is 5% in the screened population.

Test #1

Sensitivity= 80%
Specificity= 70%

Test #2

Sensitivity= 85%
Specificity= 50%

How can one compare these tests for clinical utility?

Assume 10,000 as a population sample and a prevalence of 5%:

Test #1

Sensitivity= 80%
Specificity= 70%

	Disease +	Disease -
Test +	400	2850
Test -	100	6650
	500	9500

Predictive Values

PPV: 400/3250 = 12.3%
NPV: 6650/6750 = 98.5%

Test #2

Sensitivity= 85%
Specificity= 50%

	Disease +	Disease -
Test +	425	4750
Test -	75	4750
	500	9500

PPV: 425/5175 = 8.2%
NPV: 4750/4825 = 98.4%

Which test is better?

Likelihood ratios compare probabilities of true results to false results

- Likelihood ratio of a positive test is probability of a true positive (given disease) to false positives (without disease)
- Likelihood ratio of a negative test is probability of a false negative (with disease) to a true negative (without disease)

More simply . . .

Likelihood ratio of a positive test is

$$\frac{\text{Sensitivity}}{100\% - \text{Specificity}}$$

The larger the likelihood ratio the better the ability of the test to Rule In disease

And . . .

Likelihood ratio of a negative test is

$$\frac{100\% - \text{Sensitivity}}{\text{Specificity}}$$

The smaller the likelihood ratio of a negative test the better the ability to rule out disease

Now back to Barrister's Disease

Assume 10,000 as a population sample and a prevalence of 5%:

Test #1
Sensitivity= 80%
Specificity= 70%

Likelihood ratio +: 80/100-70 = 2.66 or
Likelihood ratio -: 100-80/70 = .28 or

	Disease +	Disease -
Test +	400	2850
Test -	100	6650
	500	9500

PPV: 400/3250 = 12.3%
NPV: 6650/6750 = 98.5%

Test #2
Sensitivity= 85%
Specificity= 50%

Likelihood ratio +: 85/100-50 = 1.7
Likelihood ratio -: 100-85/ 50 = .30

	Disease +	Disease -
Test +	425	4750
Test -	75	4750
	500	9500

PPV: 425/5175 = 8.2%
NPV: 4750/4825 = 98.4%

Now in a less busy slide

Back to Barrister's Disease

Test #1		Test #2	
Sensitivity=	80%	Sensitivity=	85%
Specificity=	70%	Specificity=	50%
PPV	= 12.3%	PPV	= 8.2%
NPV	= 98.5%	NPV	= 98.4%
LR+	= 2.66	LR +	= 1.7
LR -	= .28	LR -	= .30

Making Clinical Decisions

- Statistics show you which test is more likely to yield a positive or a negative results
- What are the results to my patient
  - The costs of false positives or false negatives
  - Morbidity, Cost and Consequences

What's the Question

- In our example: is a false positive equal in clinical value to a false negative
  - If so then the likelihood ratios will tell you which is best
  - If not you are back to comparing the numbers of false positives or false negative
    - In our example if false negatives are worse then test #2 is better 75/10000 versus 100/10000 (25% relative change)
    - If false positives outweigh false negatives then test #1 is better 2850/10000 versus 4750/10000 (36% relative change)

Now that you have the tools….

- Should you screen for the disease ?
- In whom to screen ?
- How to do the screening ?
- When should screening start and how often?

Should you screen?

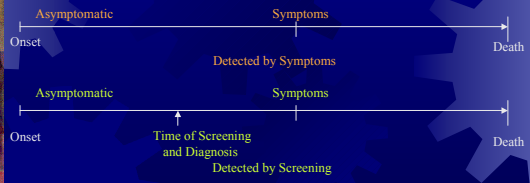
- Screening should be done if a particular disease will go on to cause substantial morbidity or mortality (what's the harm in a little football…)
- Screening is of little utility if the natural history of the disease cannot be changed or treatment of asymptomatic cases is not different from symptomatic ones.

What can go wrong?

- Diagnostic Test Errors
  - Biologic Variability
  - Measurement error
    - random
    - systematic
  - Intra observer variability
  - Inter observer variability

Lead time bias

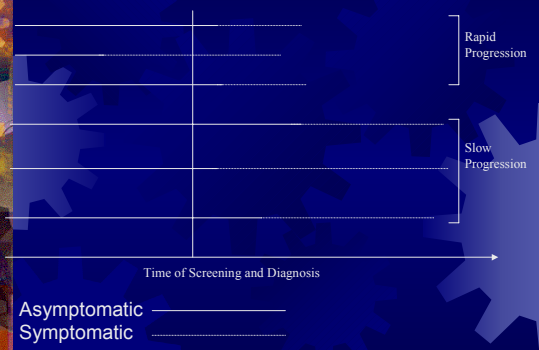
A problem found when determining if screening and subsequent treatment changes the natural history



Length bias

- Causes one to conclude that screening does not change outcomes
- In disease with a heterogeneous population, slowly progressive states will be caught more frequently by screening (prior to symptoms) and skews the data towards less benefit. Randomization usually eliminates this by sampling equal numbers of each disease subset

Length bias



Self Selection Bias

If you create the screening protocol two patient populations appear

- Those who make sure they are screened repeatedly (diminishing returns)
- Those who are rarely seen and probably most need it (i.e. TB screening yields higher benefit in homeless persons but they rarely seek medical care --- access issues)

Who do you screen?

If the disease incidence or prevalence is low what is the utility of screening

Given a Test with a sensitivity of 80% and a specificity of 90%:

	Pretest Probability			
	1%	10%	50%	90%
Positive Predictive Value	7.5%	47.1%	88.9%	98.6%
Negative Predictive Value	99.8%	97.6%	81.8%	33.3%

How do you overcome this?

By serial or consecutive testing

- Positive if Both
  - Both tests verify the same population
    - HIV testing with Elisa and Western Blot
- Positive if One
  - Two tests to detect different types of disease
    - Flexible Sigmoidoscopy and FOBT testing

Which test to use first?

- Now back to the Likelihood ratios if false positives and negative are equal
- If not then….
  - Time to be a clinician

When should you start to screen and How often?

- Natural History
- Morbidity and Mortality
- Effective Treatment given early stage

Who gets the disease?

- Screening men for ovarian cancer?
- Screening children for prostate cancer?

More seriously…..

Cholesterol in adults, Pap smears and Mammograms in adult women, colon cancer screening at 50

Does it matter?

- Does the disease cause significant morbidity and mortality
  - In economic terms is it cost effective to screen if the incidence of disease is so rare (newborn screenings for thyroid, PKU or ultrasound for ovarian cancer) or
  - Cost prohibitive to perform (AAA ultrasound screening) in a low prevalence

Can I make a difference?

- Early detection doesn't matter if treatment
  - Isn't effective
  - Not available
  - Has a high morbidity or mortality
  - Isn't readily available

So are you totally confused??

- Basically:
  - Inherent properties of the test: precision, accuracy, clinical reproducibility
  - Biologic variation: is the population operating curve well defined with little overlap between healthy and sick
  - Is the Gold Standard truly gold?
  - Test Characteristics: sensitivity, specificity and the difference to PPV,NPV
  - Is the Likelihood ratio high for Rule In or low for Rule out, and has the weight of false positives and negatives been examined

Think like a doctor??

- Again
  - Is there substantial morbidity if not treated
  - Does finding it make a clinical difference
  - Could length or lead bias explain the difference (read about Prostate and Breast Cancer now)
  - Can you improve outcomes with duel testing strategies

Thank you