SESSION #3
THERAPY

EVALUATING THE EVIDENCE
VALIDITY AND RESULTS
TODAY WE FOCUS ON VALIDITY

REMINDER: THE EBM PROCESS

- OUR PATIENT
- QUESTION (PICO)
- SEARCHING FOR AN ANSWER (EVIDENCE)
- APPRAISING THE EVIDENCE
  - EVALUATING FOR VALIDITY
  - ANALYZING THE RESULTS
- APPLYING THE RESULTS TO OUR PATIENT

THERAPY

- WE HAVE ARRIVED AT THE NEXT STEP OF THE EBM PROCESS AS NOTED IN THE PREVIOUS SLIDE – APPRAISAL OF THE EVIDENCE (THE ARTICLES OBTAINED FROM A SEARCH)
- WE WILL BE LEARNING HOW TO APPRAISE A THERAPY ARTICLE
- THERE ARE 2 ISSUES TO CONSIDER WHEN APPRAISING EVIDENCE:
  METHODOLOGIC VALIDITY AND RESULTS
- WE WILL FOCUS ON VALIDITY FOR THIS SESSION

VALIDITY FOR A THERAPY ARTICLE

THESE ARE THE BASIC VALIDITY QUESTIONS

- WAS THE ASSIGNMENT OF PATIENTS TO TREATMENT RANDOMIZED?
- WAS FOLLOW-UP SUFFICIENTLY LONG AND COMPLETE?
- WERE ALL PATIENTS ANALYZED IN THE GROUPS TO WHICH THEY WERE RANDOMIZED (INTENTION TO TREAT)?
- WERE PATIENTS AND CLINICIANS KEPT BLIND TO TREATMENT?

THESE, AS WELL AS A FEW OTHER POINTS, WILL BE DISCUSSED IN THE FOLLOWING PAGES

EBM:
Therapy

EXAMPLE

Illustrating analysis of results, especially absolute risk reduction (ARR), number needed to treat (NNT), and 95% confidence intervals

THERAPY

- 13 year old
- 1 year history of migraines
- Affecting school
- Limited relief with acute medications
- Her parents ask about using **Topiramate**

Answerable Clinical Question

- P: In patients with chronic migraine headaches,
- I: what is the therapeutic efficacy of topiramate,
- C: compared to placebo,
- O: in cutting in half the headache frequency?

- What follows are the results from a study evaluating the therapeutic benefit of Topiramate compared to placebo (it was a randomized, double blind, intention to treat study – i.e., valid)

Analyzing the Results

The Results placed in a 2X2 Table

	< 50% Reduction in headaches	≥ 50% Reduction in headaches	Total
Placebo (Control)	88	26	114
Topiramate (100mg) (Experimental)	61	59	120

Analyzing the Results

	< 50% Reduction in headaches	≥ 50% Reduction in headaches	Total
Placebo (Control)	88	26	114
Topiramate (100mg) (Experimental)	61	59	120

CER = Control Event Rate = 88/114 = 0.77 = 77%

EER = Experimental Event Rate = 61/120 = 0.51 = 51%

Analyzing the Results

	< 50% Reduction in headaches	≥ 50% Reduction in headaches	Total
Placebo (Control)	88	26	114
Topiramate (100mg) (Experimental)	61	59	120

CER = Control Event Rate = 88/114 = 0.77 = 77%

EER = Experimental Event Rate = 61/120 = 0.51 = 51%

Absolute Risk Reduction =

ARR = CER – EER

= 0.77 - 0.51 = 0.26

= 77% - 51% = 26%

Analyzing the Results

	< 50% Reduction in headaches	≥ 50% Reduction in headaches	Total
Placebo (Control)	88	26	114
Topiramate (100mg) (Experimental)	61	59	120

CER = Control Event Rate = 88/114 = 0.77 = 77%

EER = Experimental Event Rate = 61/120 = 0.51 = 51%

Absolute Risk Reduction = ARR = CER – EER = 0.77 - 0.51 = 0.26 (26%)

Number Needed to Treat

= NNT = 1/ARR = 1/0.26 = 4

95% Confidence Interval

- 95% Confidence Interval (CI)
  - If the study were repeated 100 times, 95 out of 100 times the result would be found within the 95% CI
  - You can be 95% confident that the "true" result is found within the 95% CI
- The bigger the sample, the "tighter" the 95% CI

- ARR = 26% [15%, 38%] *Statistically significant*
- NNT = 4 [3, 7]

NNT = 1/ARR = 1/0.26 = 4

Applicability – Treatment Threshold

- Are you confident "enough" that you will not have to treat more patients than your personal NNT (= treatment threshold) to see the benefit in one patient?
- Do you want to be **95% confident** that the actual number of patients you will have to treat to see the benefit in one patient is no more than your personal cutoff?
- Look at the upper end of the 95% CI

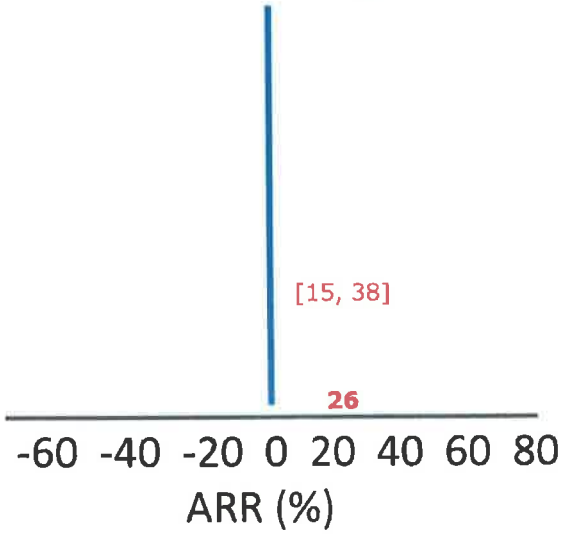
Applicability

- ARR = 26% [15%, 38%]
- NNT = 4 [3, 7]

Example: "How many patients would you be willing to treat in order for one patient to benefit from 100 mg/d topiramate (50% reduction in headache frequency)?"

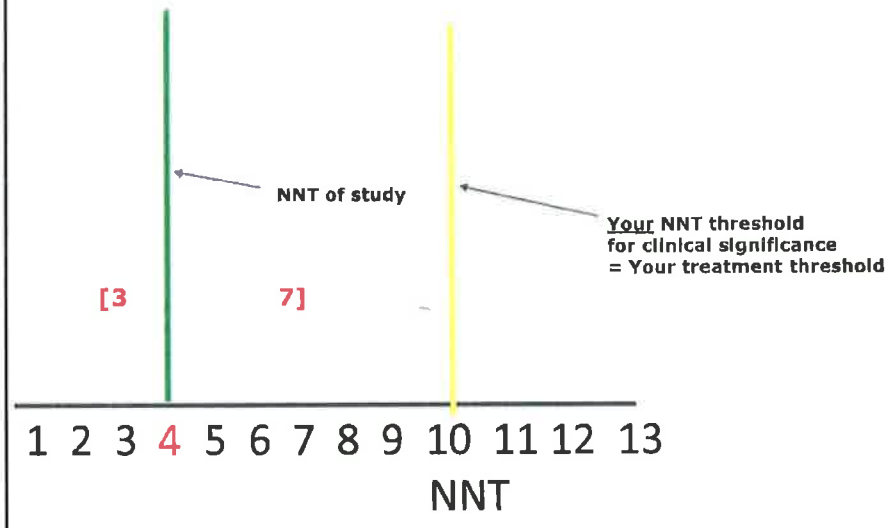
- If you want to be **95% confident** that the actual number of patients you will have to treat to see the benefit in one patient is no more than your personal cutoff...
  - **Look at the upper end of the 95% CI**

Statistical Significance



Clinical Significance

Assume your treatment threshold is NNT=10



Clinical Significance

Assume your treatment threshold is NNT=10

- "If you're willing to treat 10, you're willing to treat 7"
- i.e., you are 95% confident that the actual number of patients you will have to treat is no more than your personal treatment threshold
- Therefore, the results are both statistically and clinically significant

Evaluation of
Diagnostic Test Studies

Validity

Validity
Three Main Issues

1. Was there and independent, blind comparison to an acceptable reference (gold) standard?
2. Was the patient spectrum appropriate?
3. Was the reference (gold) standard applied regardless of the new test results?

Diagnosis: Validity

- Was there and independent, blind comparison to a "gold" or reference standard?
  - Study patients must undergo *both* tests: the new test and the reference (gold standard) test
  - The new test and "gold standard" must be assessed independently of each other by interpreters unaware of the results of the other investigation. This avoids over- or under-interpretation of the reference (gold) standard, either of which could affect study results.

Diagnosis: Validity

- Was the patient spectrum appropriate?
  - The spectrum of patients should be similar to those whom the diagnostic test will be applied in our clinical practice
  - The study patients should have varying likelihoods of having the disease. The studied patient population should not be composed of completely healthy patients (i.e., "controls") or patients that are obviously symptomatic with the disease. In both of these types of patients, testing for the disease would be unnecessary and would skew results, with the test performing better in the study population than in the typical clinical venue.
  - The spectrum of studied patients should included early and late, mild and severe cases. Also included in the spectrum of patients studied should be all common presentations of the target disorder, as well as patients with other, commonly confused diagnoses

Diagnosis: Validity

- Was the reference (gold) standard applied regardless of new test results?
  - Did the results of the new test influence the decision to perform the reference standard?
  - If so, it will lack confirmation by the "gold" standard. This could inflate the "accuracy" of the new test.
  - At times, a substitute for the gold standard may be employed when it may be unethical or impractical to use the gold standard in patients that test negative. An example of this would be a study of the diagnostic accuracy of CT scan in appendicitis. In study patients that are a lower risk for appendicitis and have a negative CT scan, one would be reluctant to perform surgery (the gold standard). A "proxy gold standard should be described in the article. In this case, long-term follow-up could be a proxy gold standard.

Diagnostic Test Studies

Understanding Results

Learning objectives:

1. Importance of pre-test probability
2. Sensitivity/Specificity
3. Likelihood Ratio (LR): MEMORIZE THE DEFINITION - it will be on the IRAT
4. The LR is pre-test-probability independent
5. Calculating the post-test probability

1

What a diagnostic test does

Pre-test probability (*Probability that the patient has disease prior to administering the test*)	x	"Results of diagnostic test"	=	Post-test probability (*Probability that the patient has disease given the additional information of the test results*)

What is a pre-test probability and where can we find it?

Pre-test probability

- Best: prevalence among my patients
- If don't know, then..
  - prevalence noted in the clinical study
  - ask a local expert
  - make an educated guess

A pre-test probability **MUST** be assigned in order to figure out the post-test probability

"Results of diagnostic test"

Sensitivity and Specificity

- Sensitivity is the proportion of people with a disease who test positive
- Specificity is the proportion of people without a disease who test negative

Example of a 2x2 table

	Truly sick	Truly well	
Positive test	TP 95	FP 100	195
Negative test	FN 5	TN 800	805
	100	900	1000

Sensitivity: How good is the test when you're sick?

	Truly sick	
Positive test	TP 95	
Negative test	FN 5	
	100	

This test correctly picks up 95/100 people who are sick.

The sensitivity is 95%

Specificity: How good is the test when you're healthy?

This test correctly classifies 800/900 people who are healthy.

Its specificity is 89%

	Truly well
Positive test	FP: 100
Negative test	TN: 800
	900

Using sensitivity/specificity

- Sensitivity and specificity are test characteristics that are independent of disease prevalence (pre-test probability)
- With sensitivity, specificity, and *your patient's* pre-test probability, you can compute *your patient's* post-test probability of having the disease
- One nice way to compute the post-test probability of disease with sensitivity and specificity is with the Likelihood Ratio

Are also Likelihood Ratios independent of disease prevalence (pre-test probability)?

YES!

(LR's are combinations of sensitivity and specificity)

Likelihood Ratio (LR)

MEMORIZE THIS AND THINK ABOUT IT!

Definition of LR:

[for any given test result]

"The probability that the patient comes from the sick rather than the healthy population"

Likelihood Ratio

MEMORIZE THIS AND THINK ABOUT IT.

For any given test result, "The probability that the patient comes from the sick rather than the well population"

- Each test result (e.g., positive, negative) has a likelihood ratio (LR+, LR-)
  - +LR should be greater than 1
  - -LR should be less than 1 (fractional)
- LR of 1 means the test result adds no new information (result is equally likely to occur in a sick as in a well person)

+LR means the LR for a positive test
-LR means the LR for a negative test

What a diagnostic test does

Pre-test "probability" (*Probability that the patient has disease prior to administering the test*)	X	Likelihood Ratio (Inherent Test Property)	=	Post-test "probability" (*Probability that the patient has disease given the additional information of the test results*)

Calculation of LR's

Notice that the LR is a combination of SENSITIVITY AND SPECIFICITY

	Truly sick	Truly well	
Positive test	TP 95	FP 100	195
Negative test	FN 5	TN 800	805
	100	900	1000

+LR = [95/100]/[100/900]
+LR = sensitivity/(1-specificity)
+ LR = 8.55

-LR = [5/100]/[800/900]
-LR = (1-sensitivity)/specificity
-LR = 0.056

+LR means the LR for a positive test
-LR means the LR for a negative test

13

The remainder of the slides discuss:

Calculating the Post-Test Probability from the Pre-Test Probability and LR

the mathematical way
the nomogram way
the online, Dr. Alan Schwartz, way

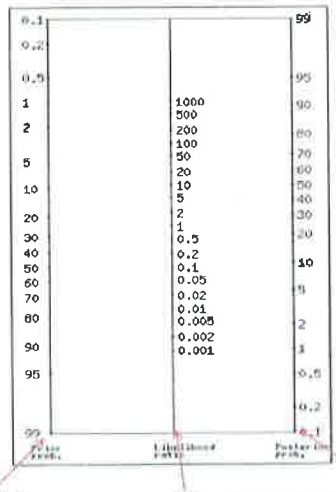
14

the mathematical way

Pre-test probability (really the odds) (*Probability that the patient has disease prior to administering the test*)	X	Likelihood Ratio (Inherent Test Property, Prevalence Independent)	=	Post-test probability (*Probability that the patient has disease given the additional information of the test results*)

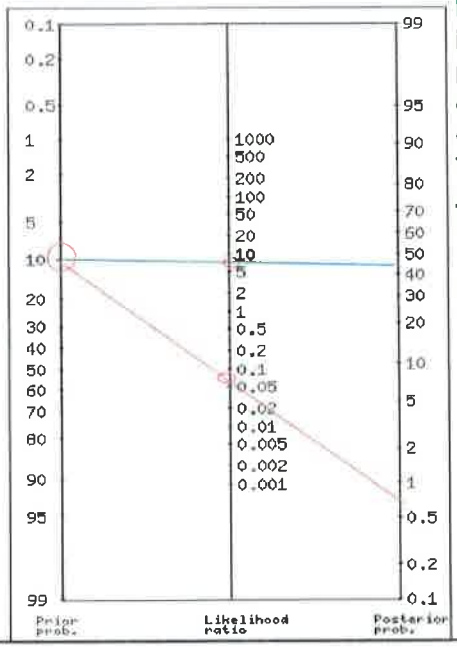
- Convert the pre-test probability (prevalence) to the pre-test odds
  pre-test odds (Pr) = prevalence/(1 – prevalence)
- Then calculate the post-test odds: **Pr x LR = post-test odds of disease**
- Finally, convert the post-test odds back to a probability
  Probability of disease = [post-test odds]/[1 + post-test odds]

the nomogram way



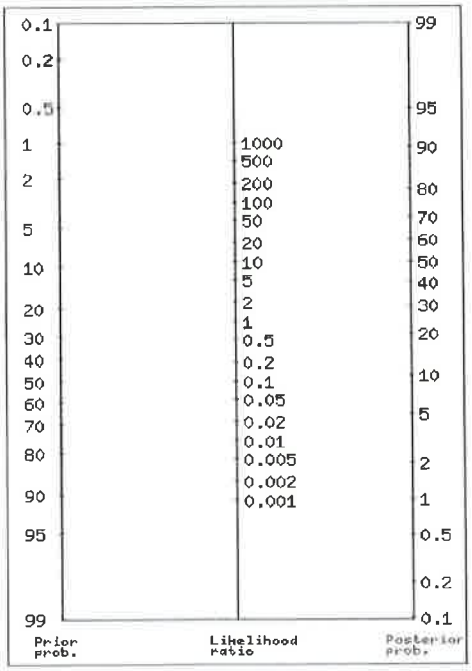
Pre Test Probability LR Post-Test Probability

Diagnostic Test Nomogram



For example, if the pre-test probability is 10% and the +LR is 9 and –LR is 0.05, then the post-test probability for a positive test is ~45% and for a negative test is ~1%

Here is spare nomogram

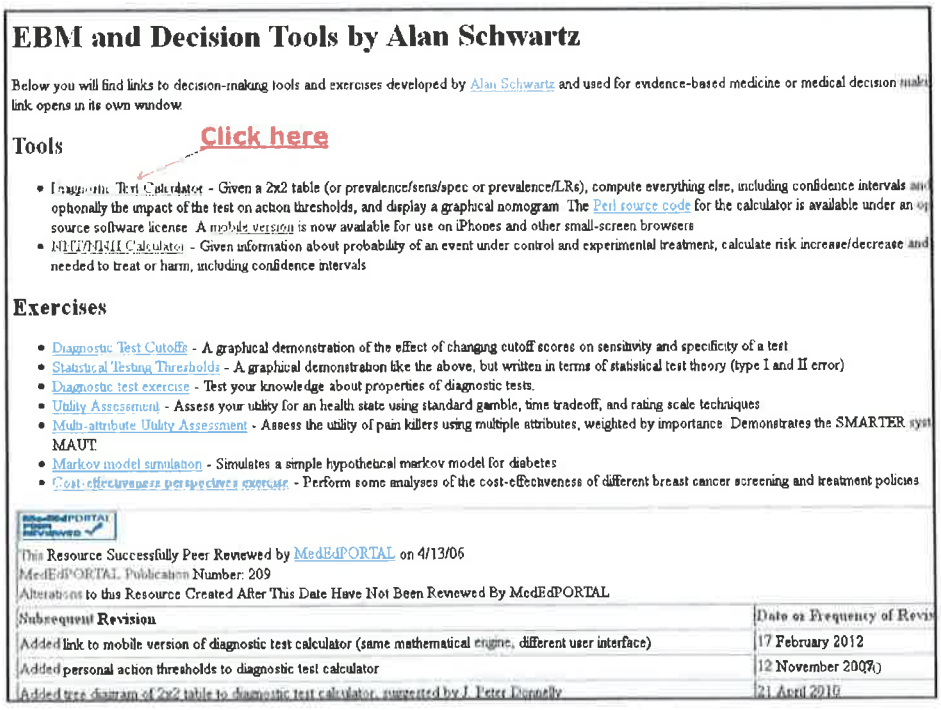


the online, Dr. Alan Schwartz, way

http://ulan.mede.uic.edu/~alansz/tools.html

OR Google EBM ALAN – First Hit

(The website will do *all* your calculations)



EBM and Decision Tools by Alan Schwartz

Below you will find links to decision-making tools and exercises developed by Alan Schwartz and used for evidence-based medicine or medical decision making. Each link opens in its own window.

Tools

Click here

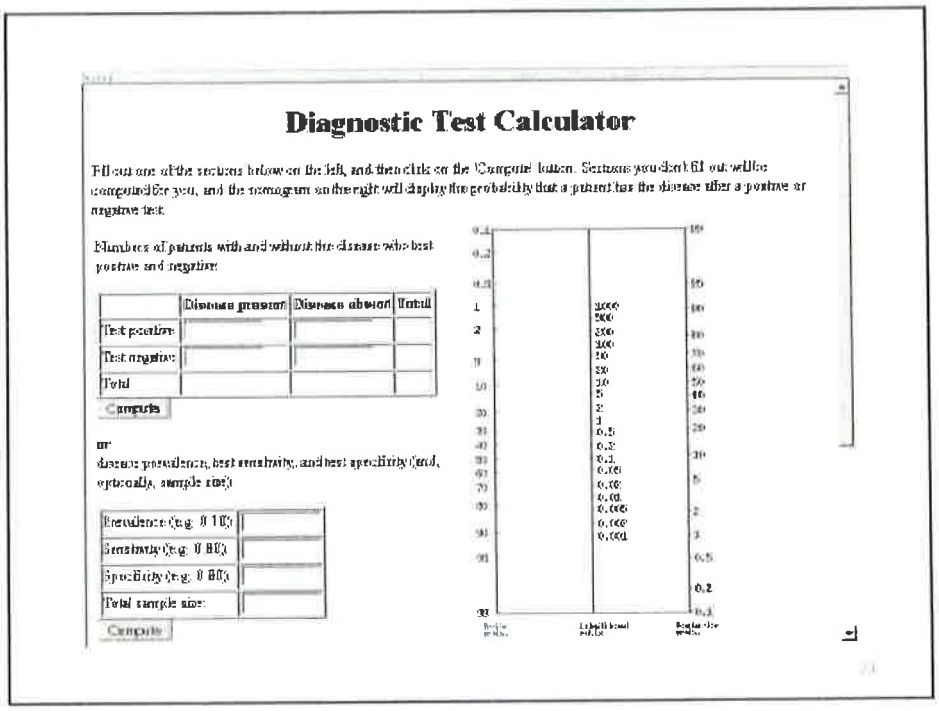
- Diagnostic Test Calculator - Given a 2x2 table (or prevalence/sens/spec or prevalence/LRs), compute everything else, including confidence intervals and optionally the impact of the test on action thresholds, and display a graphical nomogram. The Perl source code for the calculator is available under an open source software license. A mobile version is now available for use on iPhones and other small-screen browsers
- NNT/NNH Calculator - Given information about probability of an event under control and experimental treatment, calculate risk increase/decrease and number needed to treat or harm, including confidence intervals

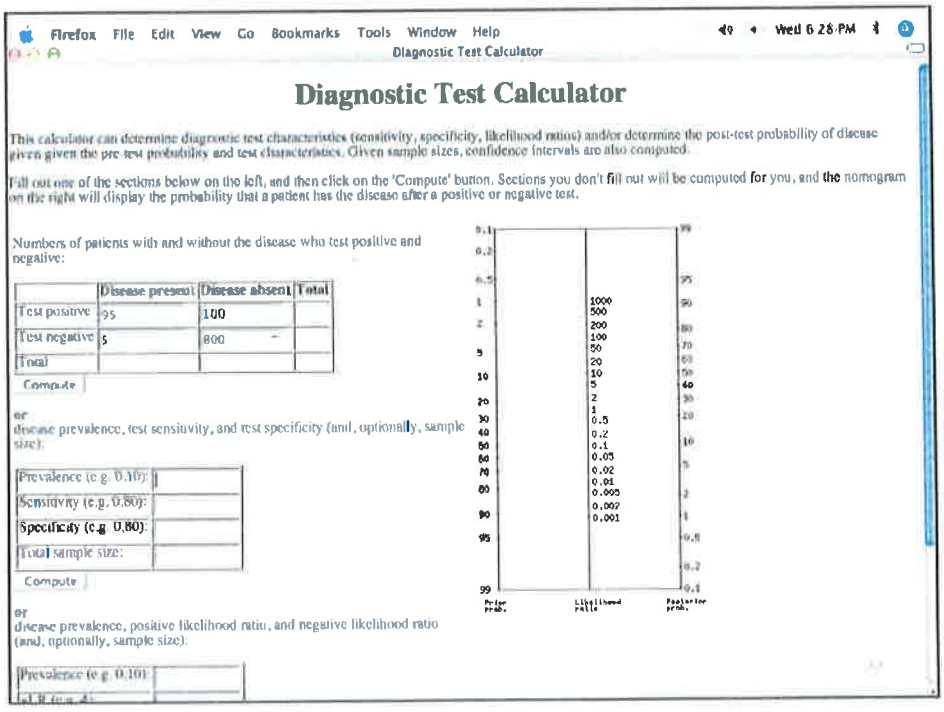
Exercises

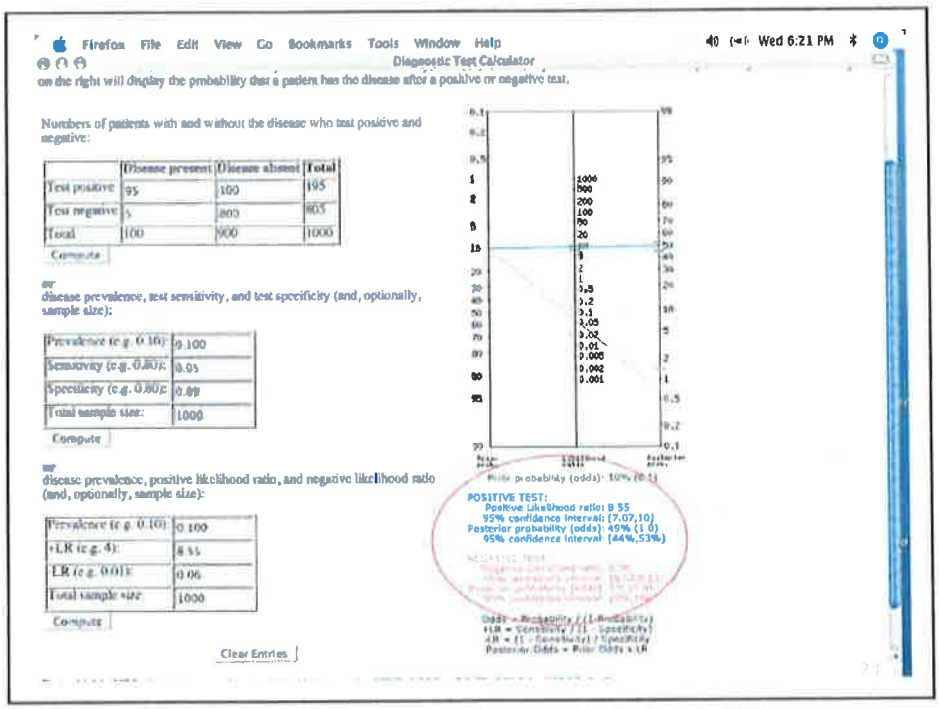
- Diagnostic Test Cutoffs - A graphical demonstration of the effect of changing cutoff scores on sensitivity and specificity of a test
- Statistical Testing Thresholds - A graphical demonstration like the above, but written in terms of statistical test theory (type I and II error)
- Diagnostic test exercise - Test your knowledge about properties of diagnostic tests.
- Utility Assessment - Assess your utility for an health state using standard gamble, time tradeoff, and rating scale techniques
- Multi-attribute Utility Assessment - Assess the utility of pain killers using multiple attributes, weighted by importance. Demonstrates the SMARTER system of MAUT
- Markov model simulation - Simulates a simple hypothetical markov model for diabetes
- Cost-effectiveness perspectives exercise - Perform some analyses of the cost-effectiveness of different breast cancer screening and treatment policies

This Resource Successfully Peer Reviewed by MedEdPORTAL on 4/13/06
MedEdPORTAL Publication Number: 209
Alterations to this Resource Created After This Date Have Not Been Reviewed By MedEdPORTAL

Subsequent Revision	Date or Frequency of Revis
Added link to mobile version of diagnostic test calculator (same mathematical engine, different user interface)	17 February 2012
Added personal action thresholds to diagnostic test calculator	12 November 2007()
Added tree diagram of 2x2 table to diagnostic test calculator, suggested by J. Peter Donnelly	21 April 2010


Diagnostic Test Calculator
Fill out one of the sections below on the left, and then click on the 'Compute' button. Sections you don't fill out will be computed for you, and the nomogram on the right will display the probability that a patient has the disease after a positive or negative test.
Numbers of patients with and without the disease who test positive and negative:
Disease present
Disease absent
Total
Test positive
Test negative
Total
Compute
or
disease prevalence, test sensitivity, and test specificity (and, optionally, sample size):
Prevalence (e.g. 0.10):
Sensitivity (e.g. 0.80):
Specificity (e.g. 0.80):
Total sample size:
Compute


Firefox File Edit View Go Bookmarks Tools Window Help
Wed 6:28 PM
Diagnostic Test Calculator
Diagnostic Test Calculator
This calculator can determine diagnostic test characteristics (sensitivity, specificity, likelihood ratios) and/or determine the post-test probability of disease given given the pre-test probability and test characteristics. Given sample sizes, confidence intervals are also computed.
Fill out one of the sections below on the left, and then click on the 'Compute' button. Sections you don't fill out will be computed for you, and the nomogram on the right will display the probability that a patient has the disease after a positive or negative test.
Numbers of patients with and without the disease who test positive and negative:
Disease present
Disease absent
Total
Test positive
95
100
Test negative
5
800
Total
Compute
or
disease prevalence, test sensitivity, and test specificity (and, optionally, sample size):
Prevalence (e.g. 0.10):
Sensitivity (e.g. 0.80):
Specificity (e.g. 0.80):
Total sample size:
Compute
or
disease prevalence, positive likelihood ratio, and negative likelihood ratio (and, optionally, sample size):
Prevalence (e.g. 0.10):


Firefox File Edit View Go Bookmarks Tools Window Help
Wed 6:21 PM
Diagnostic Test Calculator
on the right will display the probability that a patient has the disease after a positive or negative test.
Numbers of patients with and without the disease who test positive and negative:
Disease present
Disease absent
Total
Test positive
95
100
195
Test negative
5
800
805
Total
100
900
1000
Compute
or
disease prevalence, test sensitivity, and test specificity (and, optionally, sample size):
Prevalence (e.g. 0.10):
0.100
Sensitivity (e.g. 0.80):
0.95
Specificity (e.g. 0.80):
0.89
Total sample size:
1000
Compute
or
disease prevalence, positive likelihood ratio, and negative likelihood ratio (and, optionally, sample size):
Prevalence (e.g. 0.10):
0.100
+LR (e.g. 4):
8.55
-LR (e.g. 0.01):
0.06
Total sample size:
1000
Compute
Clear Entries
POSITIVE TEST:
Positive Likelihood ratio: 8.55
95% confidence interval: (7.07,10)
Posterior probability (odds): 49% (1.0)
95% confidence interval: (44%,53%)