SCREENING TESTS:
HOW TO RUIN A
PERFECTLY GOOD
MARRIAGE

David A. Grimes, M.D.
DISCLOSURES

- I have financial conflicts of interest relevant to this presentation.

- I serve on several Data Safety Monitoring Boards for clinical research sponsored by Bayer.
CONSULT, PLEASE!

- Physician’s wife gets cervical cytology, gonorrhea, and chlamydia screening at routine office visit with resident
- Gonorrhea test returns positive
- Resident has rotated to outside hospital
- You are asked to follow up with the patient
YOUR OPTIONS.....

- Tell the patient she has gonorrhea
- Repeat the test
- Get a gonorrhea culture
- Ignore the report
- None of the above
Since that time multiple papers have appeared on the use of transvaginal evaluation of ovarian masses and the characterization of the blood flow by Doppler technique. Our own recently published work has demonstrated the use of transvaginal ultrasonography in identifying very small, nonpalpable masses in asymptomatic postmenopausal women.

Because of the tremendous advances that have been made in ultrasonographic technology within the past 2 years, I believe that basing such an important conclusion on outmoded technology is a disservice not only to your readers, but also to the millions of women who are concerned about their risk for ovarian carcinoma. There is no doubt that a large multicenter study is needed to evaluate the use of modern diagnostic ultrasonography in the identification of early ovarian carcinoma. Campbell and Bourne and their colleagues have made a good start. Meanwhile it is counterproductive to deny our patients such a valuable imaging study while waiting the 5 to 10 years it will take for final results.

Barbara B. Gosink, MD

Department of Radiology, University of California, San Diego, Medical Center, 225 Dickinson St., San Diego, CA 92103-1990
Letters to the Editor

CRAMMING AGAIN

Dear Editor,

Every year parents go through this same trauma. The ruthless competition for places at university faculties is all worked out according to the final grades attained in the end-of-year examinations. Sadly, there is some cheating. However, it is often someone who lives in a small village who gets the highest grade, proving that honesty prevails for the most part. It has been recommended that some different system be introduced – but for the time being this age-old system continues to haunt parents and students alike.

Roger Daniels
Zamalek

BLOOD CLOT SCAN

Usually a physical check-up includes the heart, the liver, and a check for cancer cells – but it is rather rare to include a scan of the legs to see if there is a possibility of blood clots developing in the future. Several people I know have had strokes due to blood clots – and even cardiac arrests. After the age of 50, a blood-clot scan should surely be a must.

Merrut Mohhtan
Mohandeseen

HORSE AND CARRIAGE

A great danger to the traffic in the centre of Cairo is the carriages drawn by horses. They are parked in front of hotels to attract tourists for a general ride around Cairo. Taxis could also do the job as well. People living here no longer use these carriages, but they might give tourists the idea that we are in fact still using them when they are in front of hotels. These carriages do well in Luxor or Aswan, but not in Cairo.

Paula Winstapho
Zamalek
Heart And Vascular Screening Program
Take Control - Prevent An Event!

Your heart and blood vessels have an important job to do. Take advantage of the following screening tests:

- **Abdominal Aneurysm Scan** - looks for enlargement that could lead to rupture.
- **Carotid Artery Scan** - evaluates blood flow and looks for a risk of stroke.
- **Echocardiogram** - looks at your heart in action.
- **Lower Extremity Screening** - evaluates the arteries in your legs.
- **EKG** - picture of your heart's electrical activity.
- **Blood Pressure** - looks for hypertension.
- **Cholesterol Profile** - evaluates risk for heart disease and stroke.

All test results are reviewed and interpreted by Board Certified Heart Group Cardiologists.

Confidential results will be mailed directly to you in one week.

First series of tests for $229.00, add cholesterol screening for an additional $20.00

**Call today for an appointment**

**Dates and Locations:**

**In Ft. Myers:**
- Tuesday, February 3, 2004
- Wednesday, February 4, 2004
- Thursday, February 5, 2004
- Friday, February 6, 2004

Southwest Florida Heart Group
8540 College Parkway
Fort Myers, FL 33919
Phone: 433-8800

**In Bonita Springs:**
- Tuesday, February 10, 2004
- Thursday, February 12, 2004
- Tuesday, February 17, 2004

Southwest Florida Heart Group
Bonita Community Health Center
3501 Health Center Blvd., Ste. 2330
Bonita Springs, FL 34135
Phone: 992-9335
OBJECTIVES

- Define screening
- Define validity and its four indices
- Calculate these indices
- Describe the relationship between sensitivity and specificity
- Describe the impact of prevalence on predictive values
- Name two screening biases
SCREENING DEFINED

Testing of a large group of apparently well persons to identify those with a high probability of disease

(Pap) Diagnostic test

(Colposcopy/biopsy) Treatment

(Frostbite)
SCREENING SEMANTICS

- Can a pregnant diabetic woman be screened for hypertension?
- No, since she is not “apparently well”
- Looking for other disease: “case-finding”
CRITERIA FOR A SCREENING TEST

- Important disease
- Diagnostic and treatment facilities available
- Latent period
- Relatively prevalent disease
- Acceptable test
- Reliable (reproducible) test
- Valid test
- Appropriate to the population screened
- Cost is reasonable
Athena Gets Key Research
Duke study may lead to Alzheimer test

By Alex Barnum
Chronicle Staff Writer

A Bay Area biotechnology company has obtained rights to ground-breaking research that promises to lead to the first test that can predict whether someone will develop Alzheimer’s disease, the degenerative brain disorder that afflicts millions of Americans.
SCREENING VS. DIAGNOSIS

SCREENING
- Simple
- Low reliability
- Non-physician
- Apparently well persons

DIAGNOSIS
- Complex
- High reliability
- Physician
- Done for indications
SCREENING VS. DIAGNOSIS?

- Stool guaiac test
- Sigmoidoscopy
- Serum cholesterol
- Blood pressure
- Electronic fetal monitoring in labor  
  (TO BE REVISITED....)
MEASURING TEST PERFORMANCE

- Validity: ability of a test to distinguish between disease and health
- 4 indices widely used since the 1940’s
- Indices relate to populations, not persons
<table>
<thead>
<tr>
<th>True positive</th>
<th>False Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>False negative</th>
<th>True negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

THE WORLD IN A BOX
SENSITIVITY AND SPECIFICITY

- **Sensitivity**: the ability of a test to identify those with the disease
- **Specificity**: the ability of a test to identify those without the disease
True positive
\[ a \]
False positive
\[ b \]
False negative
\[ c \]
True negative
\[ d \]

Sensitivity = \( \frac{a}{a+c} \)
True positive

False positive

False negative

True negative

Specificity = d/(b+d)
WHAT CLINICIANS CARE ABOUT

Predictive value positive: the likelihood that a person with a positive test has the disease

and, the flip side,

Predictive value negative: the likelihood that a person with a negative test does not have the disease
MNEMONICS

Thinking vertically:
epidemiologists and lab
directors, upright citizens in the
medical community
(columns in the 2x2 table)

Thinking horizontally:
clinicians who commonly meet patients
horizontally
(rows in the 2x2 table)
<table>
<thead>
<tr>
<th></th>
<th>DISEASE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>True positive</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>False positive</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>False negative</td>
<td>c</td>
<td></td>
</tr>
<tr>
<td>True negative</td>
<td>d</td>
<td></td>
</tr>
</tbody>
</table>

**Predictive value positive** = \( \frac{a}{a+b} \)
<table>
<thead>
<tr>
<th></th>
<th>DISEASE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>False positive</td>
<td>True positive</td>
</tr>
<tr>
<td>Pos</td>
<td>a</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Neg</td>
<td>c</td>
<td>d</td>
<td></td>
</tr>
</tbody>
</table>

Predictive value negative = \( \frac{d}{c+d} \)
• Given disease, how likely is the test to be positive?
• Given health, how likely is the test to be negative?
• Given a positive test, how likely is the person to be sick?
• Given a negative test, how likely is the person to be well?
<table>
<thead>
<tr>
<th></th>
<th>WARTY DYSPLASIA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Pos PAP TEST</td>
<td></td>
</tr>
<tr>
<td>Pos</td>
<td>15 a</td>
</tr>
<tr>
<td>Neg</td>
<td>10 c</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.60</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.93</td>
</tr>
<tr>
<td>PVP</td>
<td>0.75</td>
</tr>
<tr>
<td>PVN</td>
<td>0.88</td>
</tr>
</tbody>
</table>
GAMBINO’S RULE OF THUMB

Good test: sensitivity plus specificity > 1.5

Very good test: sensitivity plus specificity > 1.8
The Pap test is a “highly accurate” test, since the false-positive rate is only 1.5%.

Agree?

Am J Obstet Gynecol (ancient)
1000 Pap smears

35 Positive

20 Positive

15 Negative

965 Negative

125 Positive

840 Negative

Colposcopy/biopsy

Hypothetical (bogus) example
DYSPLASIA

Yes | No
----|----
Pos | 20  | 15  |
Neg | 125 | 840 |

Sensitivity = 0.14
Specificity = 0.98
PVP = 0.57
PVN = 0.87
SCREENING TESTS IN SEQUENCE

Tests may complement each other

RPR

MHA-TP

Diagnosis requires both a sensitive but nonspecific reagin test, then a specific treponemal test

Diagnosis of syphilis
County Officials Seek Answers After 'Dead' Franklin Man Found Alive In Morgue

*Five Suspended After Man Left In Body Bag For Two Hours*

POSTED: 6:26 am EST January 26, 2005
UPDATED: 5:38 pm EST January 26, 2005

LOUISBURG, N.C. -- Four paramedics and a volunteer EMS are suspended with pay after a medical examiner studying a body in a morgue discovered the person was still alive.

Larry Donnell Green, 29, was removed from the morgue at the Franklin County Sheriff's Department and taken to Duke University Medical Center in Durham, where he is listed in critical condition.

Medical examiner J.B. Perdue was documenting Green's injuries to certify a cause a death when he noticed Green breathing. Green had been declared dead by paramedics at the accident scene Monday after being hit by a car driven by 36-year-old Tamuel Jackson almost two hours earlier at the U.S. 401-N.C. 39 split.

"We were making funeral arrangements, family came by, everybody thought he had passed away," said Green's brother, Steve.
'Dead' man found alive in morgue

December 30, 2005 - 8:27AM

A 95-year-old man was moved from a hospital morgue in Portugal to a care unit after he was heard coughing several hours after a doctor declared him dead, Portuguese media reported today.
Man dies after being found alive

December 28, 2003

An elderly Vietnamese man who was found alive four months ago after spending seven hours inside a drawer at a morgue has really died this time, state-controlled media reported.

Nguyen Van Quan, 73, died on Wednesday following a long illness at Nguyen Tri Phuong hospital in Ho Chi Minh City.

His death came more than four months after his daughter and son-in-law discovered him alive there when they came to take him home for his funeral, the online newspaper VnExpress.net said.

Quan was admitted to the hospital in August after complaining of a tight chest. He then experienced heart failure, his blood pressure dropped to zero and no pulse was detected. He was sent to the hospital morgue after doctors tried for 30 minutes to resuscitate him.

Quan was discovered alive seven hours later and was sent back to the intensive care unit where he stayed until he finally passed away, the report said.
DIAGNOSING DEATH, VATICAN STYLE

“The cardinal camerlengo confirms the pope’s death by calling him by his baptismal name (Karol) three times. If the pope doesn’t answer, the camerlengo says the pope is dead.”

USA Today, April 4, 2005, page 6A
PROBLEMS WITH VATICAN DEATH TEST PERFORMANCE

Test: “Are you dead yet?”

Sensitivity: 100%

Specificity: Not so hot

Predictive values: Unknown
Diagnosing Death: Where Do You Draw The Line?

Plasma Putrescine Level

Dead

Alive

(Low) (High)
No. of Persons

<table>
<thead>
<tr>
<th>Alive</th>
<th>Dead</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
</tr>
</tbody>
</table>

Plasma Putrescine mg/dl

Sensitivity 100%

Specificity 100%

No. of Persons

Plasma Putrescine mg/dl

Alive

Dead

OVERLAP
AN INVERSE RELATIONSHIP

- Sensitivity and specificity are inversely related
- Where you put the cutoff for continuous variables should reflect the impact of getting the wrong answer
PREVALENCE PROBLEMS

- The frequency of disease in the community influences the predictive values of tests, a fact not widely appreciated.
- Marriages are breaking up today because of clinicians’ naivété.
A NEW TEST FOR CHLAMYDIA

- ChlamydiaQuik™
  (Grandiose Technologies, Inc., Dismal Seepage, OH)

- A superb test by Dr. Gambino’s criteria:
  - Sensitivity = 0.95
  - Specificity = 0.95
### The Importance of Prevalence: Chlamydia (30%)

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pos</strong></td>
<td>285</td>
<td>35</td>
</tr>
<tr>
<td><strong>Neg</strong></td>
<td>15</td>
<td>665</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>300</th>
<th>700</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity = 0.95</td>
<td>PVP = 0.89</td>
<td></td>
</tr>
<tr>
<td>Specificity = 0.95</td>
<td>PVN = 0.98</td>
<td></td>
</tr>
</tbody>
</table>
Based on a positive ChlamydiaQuik™ in the health department STD clinic with a prevalence of 30%, would you be willing to prescribe azithromycin 1.0 g by mouth for her?

Her partner?
### The Importance of Prevalence: Chlamydia (5%)

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pos</td>
<td>48</td>
<td>47</td>
</tr>
<tr>
<td>Neg</td>
<td>2</td>
<td>903</td>
</tr>
</tbody>
</table>

**Test Results**

- Sensitivity = 0.95
- Specificity = 0.95
- Positive Predictive Value (PVP) = ?
- Positive Value (PVN) = 1.00
In your office, with a Chlamydia prevalence of 5%, would you be willing to prescribe azithromycin 1.0 g by mouth for her?

Her partner?
CONSULT, PLEASE!

- Physician’s wife gets cervical cytology, gonorrhea, and chlamydia screening at routine office visit with resident
- Gonorrhea test returns positive
- Resident has rotated to outside hospital
- You are asked to follow up with the patient.....
<table>
<thead>
<tr>
<th>PCR</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>97</td>
<td>198</td>
</tr>
<tr>
<td>No</td>
<td>3</td>
<td>9702</td>
</tr>
</tbody>
</table>

- Sensitivity = 0.97
- Specificity = 0.98
- PVP = ?
- PVN = ?
GONORRHEA GOOFS

With a prevalence of 1%, the predictive value positive of the PCR is about 0.33.

With a prevalence of 1 per 1,000, the predictive value positive falls to about 0.05.

Stated alternatively, 19 times out of 20 a positive PCR test (a superb test) is wrong.
Rumor of STD led to slayings, relatives say Brinton Marcell Millsap apparently shot three women and himself Friday night because of a rumor he had been exposed to a sexually transmitted disease.
AND FINALLY,

HOW DOES ALL THIS RELATE TO ELECTRONIC FETAL MONITORING?
# Futility of Electronic Fetal Monitoring

<table>
<thead>
<tr>
<th>Intrapartum Death</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pos</td>
<td>14</td>
<td>49,991</td>
</tr>
<tr>
<td>Neg</td>
<td>3</td>
<td>49,992</td>
</tr>
</tbody>
</table>

- Sensitivity = 0.85
- Specificity = 0.50
- PVP = 0.0003
- PVN = 1.00
Electronic Fetal Monitoring as a Public Health Screening Program

The Arithmetic of Failure

David A. Grimes, MD, and Jeffrey F. Petipert, MD, PhD

Electronic fetal monitoring has failed as a public health screening program. Nevertheless, most of the four million low-risk women giving birth in the United States each year continue to undergo this screening. The failure of this program should have been anticipated and thus avoided had the accepted principles of screening been considered before its introduction. All screening tests have poor positive predictive value when searching for rare conditions such as fetal death in labor or cerebral palsy. This problem is aggravated when the screening test does not have good validity as is the case with electronic fetal monitoring. Because of low-prevalence target conditions and mediocre validity, the positive predictive value of electronic fetal monitoring for fetal death in labor or cerebral palsy is near zero. Stated alternatively, almost every positive test result is wrong. To avoid such costly errors in the future, the prerequisites for any screening program must be fulfilled before the program is begun.

(Obstet Gynecol 2010;116:E397-1400)

What Is Screening?

Screening is the testing of large numbers of apparently well persons to identify those at increased risk of disease. Because screening is done among symptom-
Electronic fetal heart rate monitoring and its relationship to neonatal and infant mortality in the United States

Han-Yang Chen, MS; Suneet P. Chauhan, MD; Cande V. Ananth, PhD, MPH; Anthony M. Vintzileos, MD; Alfred Z. Abuhamad, MD

OBJECTIVE: To examine the association between electronic fetal heart rate monitoring and neonatal and infant mortality, as well as neonatal morbidity.

STUDY DESIGN: We used the United States 2004 linked birth and infant death data. Multivariable log-binomial regression models were fitted to estimate risk ratio for association between electronic fetal heart rate monitoring and mortality, while adjusting for potential confounders.

RESULTS: In 2004, 89% of singleton pregnancies had electronic fetal heart rate monitoring. Electronic fetal heart rate monitoring was associated with significantly lower infant mortality (adjusted relative risk, 0.75); this was mainly driven by the lower risk of early neonatal mortality (adjusted relative risk, 0.50). In low-risk pregnancies, electronic fetal heart rate monitoring was associated with decreased risk for Apgar scores <4 at 5 minutes (relative risk, 0.54); in high-risk pregnancies, with decreased risk of neonatal seizures (relative risk, 0.65).

CONCLUSION: In the United States, the use of electronic fetal heart rate monitoring was associated with a substantial decrease in early neonatal mortality and morbidity that lowered infant mortality.

Key words: Apgar score, electronic fetal heart rate monitoring, infant mortality, neonatal mortality, neonatal seizure

Cancer starts → Screen → Symptoms, signs → Death

Preclinical phase

Diagnosis-to-death

Screening diagnosis-to-death

Increase in longevity due to earlier diagnosis
SUMMARY

- Screening ≠ diagnosis
- Sensitivity, specificity, and predictive values measure test validity
- Remember the table shell and definitions; disease is the “top” priority
- Sensitivity and specificity are inversely related
Screening tests perform better in high-prevalence populations.

Predictive values are influenced by disease prevalence.

Lead-time and length bias occur in the absence of randomized trials.