Health Studies 315: Handouts

- Lecture Notes
- Last Year's Exam
- Article: Wagner, JM, McKinney, P, Carpenter, JL. Does this patient have appendicitis? JAMA 276(19):1589-1594
Alicia Y. Toledano, Sc.D.

- Assistant Professor,
- Department of Anesthesia and Critical Care,
- Department of Health Studies, and
- Cancer Research Center
- Donna (DACC) - 702-6378
- Penny (DHS) - 702-2453
- toledano@dacc-41.bsd.uchicago.edu
How to Use the Results of a Diagnostic Test:
Are the Results of the Study Valid?

• Jaeschke, R, Guyatt, G, Sackett, DL. JAMA 271(5):389-391
Clinical Scenario

- understanding of how to use results of diagnostic test often based on intuition, personal experience
- should know about properties of the diagnostic test
- from original literature
Literature Search

• find a study that provides information on the properties of the diagnostic test
• your practice in general
• this patient in particular
• good paper to read
• in reputable, familiar journal
• referred to by other articles
• abstract provides relevant information
Example: Rational Clinical Examination

- high morbidity (rupture), mortality associated with appendicitis
- can be significantly reduced with treatment
- routine history and physical most effective, practical diagnostic modalities
- known pathophysiology, classic signs & symptoms
- atypical location, pregnancy -> possible difficulties in diagnosis
Example, cont.

- characteristic signs & symptoms:
- poorly localized visceral pain; anorexia, nausea, vomiting; irritation of adjacent structures/peritoneum; low grade fever; peritoneal pain localized to right lower quadrant
- most consistent signs:
- guarding (voluntary), rigidity (involuntary), rebound tenderness, Rovsing sign, psoas sign, obdurator sign, rectal examination
Example, cont.

- need reproducibility of measurements
- across cases (intrarater reliability) and
- across physicians (interrater reliability)
- theory vs. practice: poor reliability depletes accuracy
- this paper: accuracy
Example: Literature Search

• marks of high quality:
• large number of patients
• valid study design
• complete reporting of data
• reviewed titles and abstracts
• will the results help you?
• compare study subjects to patients in your clinical practice
Introduction

- important to be able to assess articles about diagnostic tests
- proliferation of technology
- GOAL: efficiently assess articles, and optimally use information they provide
- three familiar questions:
  - Are the results of the study valid?
  - What are the results of the study?
  - Will the results help me in caring for my patients?
Three Questions: Valid?

• determined by study methods
• how assembled patients
• how applied test and reference
• implies reported accuracy close enough to truth for you to invest time in reading the paper
First Valid, Then Generalizable

• (First Methods, Then Results)
Sneak Preview: Friday

- Results of the Study
- accuracy
- likelihood ratio
- Caring for Patients
- generalizable
- valuable information
- comparison with other diagnostic tests
- are patients better off if this test is used?
Judging Validity

• 1. Independent, blind comparison with reference standard
• 2. Patient spectrum representative of clinical practice
• 3. Additional guides:
  • free of / minimal verification bias
  • methods described in detail
1. Comparison with Reference

- appropriate reference standard applied to every patient
- if cannot accept reference standard, study unlikely to give valid results
- test results and reference standard assessed independently
- interpretation of new test can be influenced by knowledge of reference standard result
1. Comparison with Reference: e.g.

- appropriate reference standard applied to every patient? test results and reference standard assessed independently?
- authors state that quality depends on study design, but do not mention these
- pull the articles
- ask the authors
2. Patient Spectrum

- diagnostic test should distinguish between target states that might otherwise be confused
- spectrum should resemble clinical mix
- should include borderline (difficult) cases
- in appropriate proportions
2. Patient Spectrum: e.g.

- distinguish between target states that might otherwise be confused?
- women with ob/gyne conditions: PID, ectopic pregnancy, etc.
- spectrum resemble your clinical mix?
- inclusion criteria: ED, OR; acute abdominal pain, suspected appendicitis
- % women, age range
3. Additional Guides

- Result of diagnostic test should not influence decision to obtain reference standard
- verification bias
- ethical issue
- alternative, e.g., follow up
- proportion of sample
3. Additional Guides (cont.)

• Methods should be described in detail
• patient preparation
• test performance
• analysis, interpretation of test results
3. Additional Guides: e.g.

- Result of diagnostic test influence decision to obtain reference standard?
- Need to ask authors or pull papers
- Methods described in detail?
- Unfortunately, some of the aspects evaluated are poorly defined
- Best available papers
Review: Judging Validity

• Independent, blind comparison with reference standard
• Patient spectrum representative of clinical practice
• Additional guides:
  • free of / minimal verification bias
  • methods described in detail
Results of the Study: Accuracy

- Friday:
  - ROC curve
  - prevalence
  - likelihood ratios
- Today: foundations for Friday’s material
  - 1. sensitivity, specificity
  - 2. predictive values
1. Sensitivity and Specificity

- **Sensitivity:** *hit rate*
- proportion of people with condition that test positive
- **Specificity:** *clear rate*
- proportion of people without the condition that test negative
- Based on 2 x 2 (or series of 2 x 2) table(s)
1. Sensitivity and Specificity, cont.

<table>
<thead>
<tr>
<th>Reference Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test Result</strong></td>
</tr>
<tr>
<td><strong>Test +</strong></td>
</tr>
<tr>
<td><strong>Test -</strong></td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

- **Sensitivity**: $\frac{TP}{(TP+FP)}$
- **Specificity**: $\frac{TN}{(TN+FN)}$
1. Sensitivity and Specificity, cont.

- Sensitivity and specificity are probabilities
- Sometimes easier to work with odds
- # patients with : # patients without
- If D+, odds of T+? $\text{sens/(1-sens)} = \frac{TP}{FN}$
- within D+, # of T+ : # of T-
- If D-, odds of T-? $\text{spec/(1-spec)} = \frac{TN}{FP}$
- within D-, # of T- : # of T+
2. Predictive Values

- In practice, don’t know disease status
- know test result
- Predictive Value (PV) +: proportion of people with T+ that are D+
- PV -: proportion of people with T- that are D-
2. Predictive Values, cont.

<table>
<thead>
<tr>
<th>Test Result</th>
<th>Disease +</th>
<th>Disease -</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>TP (a)</td>
<td>FP (b)</td>
<td>TP+FP</td>
</tr>
<tr>
<td>Test -</td>
<td>FN (c)</td>
<td>TN (d)</td>
<td>FN+TN</td>
</tr>
<tr>
<td>Total</td>
<td>TP+FN</td>
<td>FP+TN</td>
<td>n</td>
</tr>
</tbody>
</table>

- PV+: TP / (TP+FP)
- PV-: TN / (FN+TN)
2. Predictive Values, cont.

- again, work with odds
- If T+, odds of D+?
  \[ \frac{PV+}{1-PV+} = \frac{TP}{FP} \]
  within T+, # of D+ : # of D-
- If T-, odds of D-?
  \[ \frac{PV-}{1-PV-} = \frac{TN}{FN} \]
  within T-, # of D- : # of D+
Scenes from the Next Class

• for a test with multiple ordered outcomes, sensitivity and specificity at each cut point -> ROC curve
• predictive values are affected by prevalence
• likelihood ratios are related to all of the above