Evidence-Based Medicine
Diagnostic studies
Prognostic studies

Antoine ELIAS

Cardiology – Vascular Medicine
Clinical Research
Evidence-Based Medicine
Clinical Research

• What does it help for?
  — …………………
  — …………………
  — …………………
  — …………………
  — …………………
Clinical Research

• **Helps to answer a clinical question**
  – Evidence-based
  – Useful for clinical practice

• **Domain:**
  – « Evidence-Based Medicine »
  – « Evidence-Based Practice »
Practice of EBHC - 5 steps

• Convert information needs into answerable questions
• Efficiently locate the best evidence
• Critically appraise the evidence for validity (closeness to truth) and usefulness (applicability)
• Integrate best evidence with patient, client, population values, professional judgement and costs, & apply in practice
• Evaluation and reflection of your performance

» Adapted from Sackett, DL and Rosenberg, WMC, 1995
« Evidence-Based Medicine »
« Evidence-Based Practice »

• Steps
  – Asking answerable questions
  – Acquiring the evidence
  – Appraising the evidence
  – Applying the evidence
Asking answerable questions

• A clinical question
• A clinical research question
Asking answerable questions

• What type of questions are met in clinical practice?
  — .................
  — .................
  — .................
  — .................
  — .................
  — .................
Asking answerable questions

- Questions raised in clinical practice
  - Diagnosis
  - Prognosis
  - Therapy
  - Etiology
  - Qualitative
  - Economic
  - Clinical Practice Guidelines
  - Systematic Review – Meta analysis
  - ............
Asking answerable questions

• **Type of question**
  – Diagnosis/Screening
  – Pronosis
  – Etiology
  – Therapy (effectiveness - safety)
  – ….

• **PI(E)COT format**
  – Population
  – Intervention (Exposure)
  – Control
  – Outcome
  – Time
Asking answerable questions

• **Study Design**
  – Observational versus Experimental
  – Cross-sectional versus Longitudinal
  – Case-Control versus Cohort
  – Prospective versus Rétrospective
  
  – (Schéma)

• **Clinical questions**
  – Diagnostic
  – Pronostic
  – Etiological
  – Therapeutic (benefit, harm)
  – Qualitative
  – Economic
  – Clinical Practice Guidelines
  – Systematic Review – Meta analysis
D GRIMES and K F SCHULZ
An overview of clinical research: the lay of the land
Lancet 2002; 359: 57–61
D GRIMES and K F SCHULZ
An overview of clinical research: the lay of the land
Lancet 2002; 359: 57–61
Asking answerable questions

Type of studies relevant to clinical research questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Etiology</td>
<td>Case-Control - Cohort</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Cross-sectional - Cohort</td>
</tr>
<tr>
<td>Pronostic</td>
<td>Inception cohort</td>
</tr>
<tr>
<td>Treatment</td>
<td>Randomised Controlled Trial (RCT)</td>
</tr>
<tr>
<td>Cost</td>
<td>Cost-Effectiveness Analysis</td>
</tr>
<tr>
<td></td>
<td>Cost-Utility Analysis</td>
</tr>
<tr>
<td>Qualitative</td>
<td>Qualitative</td>
</tr>
</tbody>
</table>
Finding the evidence

- Database
- Key-words
Finding the evidence

• Terms:
  – Text words
  – MeSH (Medical Subject Headings): vocabulaire contrôlé des mots clés dans PubMed et Cochrane

• Notation autre
  – Truncation (* ou $): child* ou child$ pour child ou children
  – Wildcard (?): gyn?ecology pour gynaecology ou gynecology

• Boolean: OR - AND

• Limits: MeSH, type of publication, year of publication

• Databases
  – PubMed (+ clinical queries)
  – EMBASE
  – Cochrane
  – NLH (National Library of Health)
  – Google
  – CINAHL (Cumulative Index to Nursing and Allied Health Literature)
  – ..... 

• Save your references (EndNote, RefManager, RefWorks)
## Finding the evidence

### Boolean Links

<table>
<thead>
<tr>
<th>Primary term</th>
<th>Synonym 1</th>
<th>Synonym 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>OR</td>
<td>OR</td>
</tr>
<tr>
<td>I</td>
<td>OR</td>
<td>OR</td>
</tr>
<tr>
<td>C</td>
<td>OR</td>
<td>OR</td>
</tr>
<tr>
<td>O</td>
<td>OR</td>
<td>OR</td>
</tr>
</tbody>
</table>
Appraising the evidence

• Internal validity
• Results
• External validity
Appraising the evidence

• **Internal validity (bias)**
  – **Representativeness** (random selection, random allocation)
  – **Ascertainment** (Response-rate, follow-up)
  – **Measurement** (*blinded/objective*)

• **Methods**
  – Descriptive
  – Study quality
GATE Frame: 5 PI(E)COT components

1. Participants
2. Exposure Gp
3. Comparison
4. Outcomes
   +  -
5. Time

GATE a general appraisal tool for epidemiology
http://www.health.auckland.ac.nz/comhealth/epiq/epiq.htm
Appraising the evidence

• What are the results? Are the results important?

• **Study impact (Effect size):**
  – Statistically significant (confidence interval, p-value)
  – Clinically significant: effect measure
    • Relative risk: RR, RRR, OR
    • Absolute risk: ARR, NNT
D GRIMES and K F SCHULZ
An overview of clinical research: the lay of the land
Lancet 2002; 359: 57–61
Appraising the evidence

• External validity
  – Utility
  – Consistency
  – Representativeness (« generalisability »)
  – Transferability and applicability
Appraising the evidence

• External validity

• – Integrate:
  – Professional judgment (clinical expertise)
  – Patient preference (value)
  – Cost
Appraising the evidence
Applying the evidence

• Integrate clinical expertise and patient values
  – Are your patients similar to those of the study? (« How different would my patient have to be for the results of the study to be of no help » - D Sackett)
  – Does the comparison intervention reflect your current practice?
  – How much of the study effect can I expect for your patient(s)?
  – Is the intervention realistic in your setting?
    • Availability (equipment, personal, local expertise)
    • Workload
    • Cost
  – What alternatives are available?
  – Are the outcomes appropriate to your patient(s)?
  – Does the intervention meet their values and preferences?

• EBM Toolkit 2nd Edition Carl Heneghan & Douglas Badenoch BMJ Books
References

• Lancet book
• EBM Sharon Strauss et al
• EBM Toolkit - Blackwell Publishing BMJ Books
• Statistics Toolkit - Blackwell Publishing BMJ Books
Diagnostic Studies
Think about diagnosis
Diagnosis

• Decision Analysis
  – Probability of Disease
    • Rule out
    • Uncertain
    • Rule in
  – Risk of Outcome
    • Recurrence and death if no treatment
    • Complications related to diagnostic methods
    • Risk of treatment
  – Cost
  – Decision Making
    • No treatment
    • Perform another test
    • Treat
Diagnostic tests
Research Methodology

• Different steps required for diagnostic studies
  – Be familiar with the test
    • Technique - Criteria – Limitations
  – Assess **Performance**
    • Reliability: consistency - reproducibility
    • **Accuracy : efficacy** (versus standard)
    • Validity: different settings
  – Assess **impact: effectiveness**
    (either as a single test or integrated within a diagnostic strategy)
    • Diagnostic management studies
    • Randomised Controlled Trials
  – **Economic evaluation: efficiency**
    • Cost-effectiveness/ Cost-utility analyses
Agreement in Measurements and Reliability studies

• Categorical variables
  – % agreement optimistic: ignores agreement by chance
  – Inter-rater agreement
    • Kappa = (% observed - % expected)/(1 - % expected)
    • Value of Kappa [95% CI]
      – 0.6-0.8: good
      – 0.8-1.0: very good

• Continuous variables
  – Correlation is inappropriate
  – Bland-Altman : (mean difference ± 2 SD) [95% CI]
  – Intraclass correlation coefficient
Diagnostic Test Accuracy

- Population
- Intervention: Index test
- Comparison: Reference test
- Outcome (accuracy)
Diagnostic Test Accuracy

- Study design
  - Common
    - All participants
    - Index test
    - Reference test
    - Results (accuracy measurements)
  - Other designs
    - More than one index test
    - Two reference tests for ethical reason (invasive test: confirmation if +ve, FU if –ve) with different time-frames
## Diagnostic Test Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th>No Disease</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test +ve</strong></td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td><strong>Test -ve</strong></td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td></td>
<td>a+c</td>
<td>b+d</td>
<td>a+b+c+d</td>
</tr>
</tbody>
</table>

Paired measures of test accuracy

- Sensitivity and Specificity
- NPV and PPV
- LR+ve and LR-ve
## Diagnostic Test Accuracy

Results of a systematic review of serum ferritin as a diagnostic test for iron deficiency anemia.

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th>No Disease</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +ve</td>
<td>731</td>
<td>270</td>
<td>1001</td>
</tr>
<tr>
<td>Test -ve</td>
<td>78</td>
<td>1500</td>
<td>1578</td>
</tr>
<tr>
<td></td>
<td>809</td>
<td>1770</td>
<td>2579</td>
</tr>
</tbody>
</table>

Sensitivity  
Specificity  
NPV  
PPV  
LR+ve  
LR-ve

[95% Confidence Interval]
# Multiple-level Test

<table>
<thead>
<tr>
<th>Levels</th>
<th>Disease</th>
<th>No Disease</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 6</td>
<td>a</td>
<td>b</td>
<td>a+b</td>
</tr>
<tr>
<td>Level 5</td>
<td>c</td>
<td>d</td>
<td>c+d</td>
</tr>
<tr>
<td>Level 4</td>
<td>e</td>
<td>f</td>
<td>e+f</td>
</tr>
<tr>
<td>Level 3</td>
<td>g</td>
<td>h</td>
<td>g+h</td>
</tr>
<tr>
<td>Level 2</td>
<td>i</td>
<td>j</td>
<td>i+j</td>
</tr>
<tr>
<td>Level 1</td>
<td>k</td>
<td>l</td>
<td>k+l</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>a+c+e+g+i+k</td>
<td>b+d+f+h+j+l</td>
<td></td>
</tr>
</tbody>
</table>

**Levels** (top highest, bottom lowest)

**Types:** risk-group (high, moderate, low), test-probability (high, moderate, low), cut-off (ex: D-dimer)
### V/Q Scan Result

<table>
<thead>
<tr>
<th>V/Q Scan Result</th>
<th>PE Present</th>
<th>PE Absent</th>
<th>P(VQi/PE+)</th>
<th>P(VQi/PE-)</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>102</td>
<td>14</td>
<td>0.40</td>
<td>0.03</td>
<td>14</td>
</tr>
<tr>
<td>Intermediate</td>
<td>105</td>
<td>217</td>
<td>0.42</td>
<td>0.45</td>
<td>0.9</td>
</tr>
<tr>
<td>Low</td>
<td>39</td>
<td>199</td>
<td>0.16</td>
<td>0.42</td>
<td>0.4</td>
</tr>
<tr>
<td>Normal / near Normal</td>
<td>5</td>
<td>50</td>
<td>0.02</td>
<td>0.10</td>
<td>0.2</td>
</tr>
</tbody>
</table>

**PIOPED JAMA 1990**
Distribution of B-type natriuretic peptide results in non-diseased (normal LVEF) and diseased (reduced LVEF) groups; two cutoffs are shown for 20 and 40 pmol/L.
<table>
<thead>
<tr>
<th>From Pre-Test probability to Post-Test Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Test Probability (Clinical Probability)</strong>: (=p)</td>
</tr>
<tr>
<td><strong>Pre-Test Odds</strong>: (= p/(1-p))</td>
</tr>
<tr>
<td><strong>Likelihood Ratio</strong>: (LR(+)\ or (LR(-))</td>
</tr>
<tr>
<td>(LR (+)) (= \frac{Se}{1-Sp})</td>
</tr>
<tr>
<td>(LR (-)) (= \frac{(1-Se)}{Sp})</td>
</tr>
<tr>
<td><strong>Post-Test Odds</strong>: (= LR \times \text{Pre-test Odds})</td>
</tr>
<tr>
<td><strong>Post-Test Probability</strong>: (= \frac{\text{Post-test odds}}{\text{Post-test odds} + 1})</td>
</tr>
</tbody>
</table>
Likelihood ratio nomogram.
Test-treatment thresholds.
### ROC curve

http://www.medicalbiostatistics.com/

<table>
<thead>
<tr>
<th>Duration (hr) greater than or equal to</th>
<th>Sensitivity</th>
<th>1 - specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>.00</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>.63</td>
<td>1.000</td>
<td>.976</td>
</tr>
<tr>
<td>.88</td>
<td>1.000</td>
<td>.969</td>
</tr>
<tr>
<td>1.25</td>
<td>1.000</td>
<td>.890</td>
</tr>
<tr>
<td>1.75</td>
<td>1.000</td>
<td>.866</td>
</tr>
<tr>
<td>2.13</td>
<td>1.000</td>
<td>.819</td>
</tr>
<tr>
<td>2.38</td>
<td>1.000</td>
<td>.811</td>
</tr>
<tr>
<td>2.75</td>
<td>1.000</td>
<td>.780</td>
</tr>
<tr>
<td>3.25</td>
<td>1.000</td>
<td>.717</td>
</tr>
<tr>
<td>3.75</td>
<td>1.000</td>
<td>.709</td>
</tr>
<tr>
<td>4.50</td>
<td>1.000</td>
<td>.646</td>
</tr>
<tr>
<td>5.13</td>
<td>.971</td>
<td>.583</td>
</tr>
<tr>
<td>5.38</td>
<td>.971</td>
<td>.575</td>
</tr>
<tr>
<td>5.75</td>
<td>.971</td>
<td>.551</td>
</tr>
<tr>
<td>6.25</td>
<td>.914</td>
<td>.378</td>
</tr>
<tr>
<td>6.75</td>
<td>.914</td>
<td>.346</td>
</tr>
<tr>
<td>7.13</td>
<td>.857</td>
<td>.291</td>
</tr>
<tr>
<td>7.38</td>
<td>.857</td>
<td>.283</td>
</tr>
<tr>
<td>7.75</td>
<td>.857</td>
<td>.276</td>
</tr>
<tr>
<td>8.25</td>
<td>.800</td>
<td>.189</td>
</tr>
<tr>
<td>8.75</td>
<td>.800</td>
<td>.181</td>
</tr>
<tr>
<td>9.25</td>
<td>.743</td>
<td>.110</td>
</tr>
<tr>
<td>9.75</td>
<td>.743</td>
<td>.102</td>
</tr>
<tr>
<td>10.25</td>
<td>.543</td>
<td>.039</td>
</tr>
<tr>
<td>10.75</td>
<td>.543</td>
<td>.031</td>
</tr>
<tr>
<td>11.50</td>
<td>.457</td>
<td>.024</td>
</tr>
<tr>
<td>12.50</td>
<td>.400</td>
<td>.008</td>
</tr>
<tr>
<td>13.50</td>
<td>.343</td>
<td>.000</td>
</tr>
<tr>
<td>14.50</td>
<td>.286</td>
<td>.000</td>
</tr>
<tr>
<td>15.50</td>
<td>.257</td>
<td>.000</td>
</tr>
<tr>
<td>16.50</td>
<td>.200</td>
<td>.000</td>
</tr>
<tr>
<td>17.50</td>
<td>.171</td>
<td>.000</td>
</tr>
<tr>
<td>18.50</td>
<td>.143</td>
<td>.000</td>
</tr>
<tr>
<td>19.50</td>
<td>.114</td>
<td>.000</td>
</tr>
<tr>
<td>20.25</td>
<td>.057</td>
<td>.000</td>
</tr>
</tbody>
</table>
ROC curve

• Select cut-off for patient management

• Cut-off selection is a clinical choice
  – Best Se?
  – Best Sp?
  – Both?

• $\frac{Se}{1-Sp} = LR+ve$

• Between 2 cut-offs:
  – $\frac{Se}{1-Sp} = \text{slope (tangent)}$

• Diagonale:
  – $Se = 1-Sp$ and $LR +ve = 1$ (non informative)
ROC AUC

- ROC AUC [95% CI]
  - Best: AUC near 1
  - Non informative (non significant if CI includes 0.5)
  - Discrimination power (helps comparing two/three/or more tests by AU ROC [95% CI])

- Disadvantage
  - Cannot be interpreted in terms of individual patients
Critical Appraisal – Diagnostic studies

• Is this evidence about a diagnostic test valid?
• Are the results important?

Does this (valid) evidence demonstrate an important ability of this test to distinguish patients who do and who do not have a specific disorder?

• Are the results of this diagnostic study applicable to my patient?
**Population:** appropriate spectrum, consecutive patients  
**Index test:** tech description, pre-established criteria  
**Reference test:** tech description, pre-established criteria  
**Patients had both tests:** the reference test regardless of the index test  
**Measurement:** independent observers – blind interpretation

---

**Table 5.2** Is this evidence about a diagnostic test valid?

1. **Representative:** Was the diagnostic test evaluated in an appropriate spectrum of patients (those in whom we would use it in practice)?
2. **Ascertainment:** Was the reference standard ascertained regardless of the diagnostic test result?
3. **Measurement:** Was there an independent, blind comparison with a reference standard?  
   (Fourth question to be considered for clusters of tests of clinical prediction rules)  
4. Was the cluster of tests validated in a second, independent group of patients?
Assessing a Study of a Test
(Jaeschke et al, JAMA, 1994, 271: 389-91)

• Was an appropriate spectrum of patients included?
  – (Spectrum Bias)
• All patients subjected to a Gold Standard?
  – (Verification Bias)
• Was there an independent, "blind" comparison with a Gold Standard?
  – Observer Bias; Differential Reference Bias
• Methods described so you could repeat test?
Are the results important?
Does this (valid) evidence demonstrate an important ability of this test to distinguish patients who do and who do not have a specific disorder?

- Reliability and efficacy performances
- Effectiveness (management studies – RCT)
Are the results of this diagnostic study applicable to my patient?

<table>
<thead>
<tr>
<th>Table 5.5</th>
<th>Are the results of this diagnostic study applicable to my patient?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Is the diagnostic test available, affordable, accurate, and precise in our setting?</td>
</tr>
<tr>
<td>2.</td>
<td>Can we generate a clinically sensible estimate of our patient’s pre-test probability?</td>
</tr>
<tr>
<td></td>
<td>a. From personal experience, prevalence statistics, practice databases, or primary studies</td>
</tr>
<tr>
<td></td>
<td>b. Are the study patients similar to our own?</td>
</tr>
<tr>
<td></td>
<td>c. Is it unlikely that the disease possibilities or probabilities have changed since this evidence was gathered?</td>
</tr>
<tr>
<td>3.</td>
<td>Will the resulting post-test probabilities affect our management and help our patient?</td>
</tr>
<tr>
<td></td>
<td>a. Could it move us across a test–treatment threshold?</td>
</tr>
<tr>
<td></td>
<td>b. Would our patient be a willing partner in carrying it out?</td>
</tr>
<tr>
<td></td>
<td>c. Would the consequences of the test help our patient reach his or her goals in all this?</td>
</tr>
</tbody>
</table>

© Straus, Glasziou, Richardson, Haynes: Evidence-Based Medicine, 4th Edition.
Critical Appraisal Tools (CAT)

- **CASP**
  - Critical Appraisal Skills Programme (CASP)

- **STARD**
  - Checklist of 25 items
  - Flowchart
  - **April 2008** - More than 200 biomedical journals encourage the use of the STARD statement in their instructions for authors.
Systematic reviews of diagnostic accuracy tests

- Cochrane Collaboration DTA Handbook
  - [http://srdta.cochrane.org/welcome](http://srdta.cochrane.org/welcome)
  - 5 chapters of Handbook are published online
  - RevMan

- QUADAS – QUALity of Diagnostic Accuracy Studies
  - [www.quadas.org](http://www.quadas.org)
  - Risk of bias tool
Examples of clinical prediction rules.

<table>
<thead>
<tr>
<th>Prediction rule</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ottawa Ankle Rule(^8)</td>
<td>99.6%</td>
<td>48%</td>
</tr>
<tr>
<td>ABCD rule for melanoma(^9)</td>
<td>84%</td>
<td>56%</td>
</tr>
<tr>
<td>Well’s DVT Rule(^10)</td>
<td>Multilevel test</td>
<td></td>
</tr>
<tr>
<td>ABCD rule for stroke prediction after TIA(^11)</td>
<td>Multilevel test</td>
<td></td>
</tr>
</tbody>
</table>

© Straus, Glasziou, Richardson, Haynes: Evidence-Based Medicine, 4th Edition.
Simplified Clinical Model for Assessment of Deep Vein Thrombosis*

Table 2. Simplified Clinical Model for Assessment of Deep Vein Thrombosis*

<table>
<thead>
<tr>
<th>Clinical Variable</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active cancer (treatment ongoing or within previous 6 months or palliative)</td>
<td>1</td>
</tr>
<tr>
<td>Paralysis, paresis, or recent plaster immobilization of the lower extremities</td>
<td>1</td>
</tr>
<tr>
<td>Recently bedridden for 3 days or more, or major surgery within the previous 12 weeks requiring general or regional anesthesia</td>
<td>1</td>
</tr>
<tr>
<td>Localized tenderness along the distribution of the deep venous system</td>
<td>1</td>
</tr>
<tr>
<td>Entire leg swelling</td>
<td>1</td>
</tr>
<tr>
<td>Calf swelling at least 3 cm larger than that on the asymptomatic leg (measured 10 cm below the tibial tuberosity)↑</td>
<td>1</td>
</tr>
<tr>
<td>Pitting edema confined to the symptomatic leg</td>
<td>1</td>
</tr>
<tr>
<td>Collateral superficial veins (nonvaricose)</td>
<td>1</td>
</tr>
<tr>
<td>Previously documented DVT</td>
<td>1</td>
</tr>
<tr>
<td>Alternative diagnosis at least as likely as DVT</td>
<td>-2</td>
</tr>
</tbody>
</table>

Abbreviation: DVT, deep vein thrombosis.
*Scoring method indicates high probability if score is 3 or more; moderate if score is 1 or 2; and low if score is 0 or less.
↑ In patients with symptoms in both legs, the more symptomatic leg was used.

Guides for critically appraising a report about pre-test probabilities of disease.

<table>
<thead>
<tr>
<th>Table 5.6</th>
<th>Guides for critically appraising a report about pre-test probabilities of disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Is this evidence about pre-test probability valid?</td>
</tr>
<tr>
<td></td>
<td>a. Did the study patients represent the full spectrum of those who present</td>
</tr>
<tr>
<td></td>
<td>with this clinical problem?</td>
</tr>
<tr>
<td></td>
<td>b. Were the criteria for each final diagnosis explicit and credible?</td>
</tr>
<tr>
<td></td>
<td>c. Was the diagnostic work-up comprehensive and consistently applied?</td>
</tr>
<tr>
<td></td>
<td>d. For initially undiagnosed patients, was follow-up sufficiently long and</td>
</tr>
<tr>
<td></td>
<td>complete?</td>
</tr>
<tr>
<td>2.</td>
<td>Is this evidence about pre-test probability important?</td>
</tr>
<tr>
<td></td>
<td>a. What were the diagnoses and their probabilities?</td>
</tr>
<tr>
<td></td>
<td>b. How precise were these estimates of disease probability?</td>
</tr>
</tbody>
</table>

© Straus, Glasziou, Richardson, Haynes: Evidence-Based Medicine, 4th Edition.
Prognostic studies
Prognosis

• Questions

• Q1. What is Prognosis?

• Q2. Why is it important to assess prognosis?

• Q3. Examples on prognosis
What is Prognosis?

• Course of disease:
  – Disease → Outcome
  – Incidence of outcome
    • death,
    • disease progression,
    • quality of life,
    • ....
Prognosis

• Predictor finding
  – Causality (etiology)

• Prediction
  – Estimation (at inception)
  – Probability/risk of an individual developing an event
  – Period of time
  – Based on individual characteristics

• Prognostic models: combination of predictors ($\geq 2$ variables)
  • Prediction models: risk stratification useful for risk prediction in individuals
  • Predictive models: ability to modify response to treatment (may be only 1 variable in model)
Why is it important to assess prognosis?

• Stakeholders
  – health care professionals (public health, doctors, health care providers, policy decision makers)
  – Individuals

• Why important?
  – To inform patient
  – To make a decision on treatment
Prognostic models

• Prognostic models – Prediction rules – Prognostic scores
  – score,
  – formulae,
  – index,
  – nomograms, ...

• SCORE (CV risk)
• Ottawa ankle rule (ankle or foot pain - X ray?)
• PESI (Pulmonary embolism severity index)
• Simplified PESI
• HERDOO2
• CHADS$_2$ -> CHA$_2$DS$_2$VASc
sPESI (simplified PESI)
Jimenez, Aubresky et al 2010 (Arch Intern Med)

Table 1. Original and Simplified Pulmonary Embolism Severity Index (PESI)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Original PESI</th>
<th>Simplified PESI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt;80 y</td>
<td>Age in years</td>
<td>1</td>
</tr>
<tr>
<td>Male sex</td>
<td>+10</td>
<td></td>
</tr>
<tr>
<td>History of cancer</td>
<td>+30</td>
<td>1</td>
</tr>
<tr>
<td>History of heart failure</td>
<td>+10</td>
<td></td>
</tr>
<tr>
<td>History of chronic lung disease</td>
<td>+10</td>
<td></td>
</tr>
<tr>
<td>Pulse ≥110 beats/min</td>
<td>+20</td>
<td>1</td>
</tr>
<tr>
<td>Systolic blood pressure &lt;100 mm Hg</td>
<td>+30</td>
<td>1</td>
</tr>
<tr>
<td>Respiratory rate ≥30 breaths/min</td>
<td>+20</td>
<td></td>
</tr>
<tr>
<td>Temperature &lt;36°C</td>
<td>+20</td>
<td></td>
</tr>
<tr>
<td>Altered mental status</td>
<td>+60</td>
<td></td>
</tr>
<tr>
<td>Arterial oxyhemoglobin saturation</td>
<td>+20</td>
<td>1</td>
</tr>
<tr>
<td>level &lt;90%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:
- A total point score for a given patient is obtained by summing the patient's age in years and the points for each predictor when present.
- The score corresponds to the following risk classes: 65 or less, class I; 66 to 85, class II; 86 to 105, class III; 106 to 125, class IV; and more than 125, class V. Patients in risk classes I and II are defined as being at low risk.
- A total point score for a given patient is obtained by summing the points. The score corresponds to the following risk classes: 0, low risk; 1 or more, high risk. Empty cells indicate that the variable was not included.
- The variables were combined into a single category of chronic cardiopulmonary disease.
Prognostic models: VTE recurrence

- Eichinger, S., et al.,
  
  *Risk assessment of recurrence in patients with unprovoked deep vein thrombosis or pulmonary embolism: the Vienna prediction model.*

- Rodger, M.A., et al.,
  
  *Identifying unprovoked thromboembolism patients at low risk for recurrence who can discontinue anticoagulant therapy.*
Men and HERDOO 2

<table>
<thead>
<tr>
<th>Box 1: Clinical decision rule* to identify women at low risk of recurrent venous thromboembolism after 5–7 months of oral anticoagulant therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women with 0 or 1 of the following features may be able to safely discontinue therapy:</td>
</tr>
<tr>
<td>- Post-thrombotic signs (hyperpigmentation, edema or redness in either leg)</td>
</tr>
<tr>
<td>- D-Dimer level ≥ 250 μg/L</td>
</tr>
<tr>
<td>- Body mass index ≥ 30 kg/m²</td>
</tr>
<tr>
<td>- Age ≥ 65 yr</td>
</tr>
</tbody>
</table>

*Generated using the following logistic regression model: \[ Y = B_1 X_1 + B_2 X_2 + \ldots; \] recurrent venous thromboembolism = \(-3.9717 \times \text{intercept} + 1.2977 \times \text{body mass index} \geq 30 \text{ kg/m}^2 \)+ 0.6473 \times \text{(either leg hyperpigmentation, edema or redness)} + 0.9155 \times \text{(D-dimer} \geq 250 \mu g/L \)+ 0.8084 \times \text{(age} \geq 65 \text{ yr).}
SCORE Chart for Use in High-Risk European Regions

CVD = cardiovascular disease; SCORE = Systematic COronary Risk Evaluation.
## CHADS₂ - CHA₂DS₂-VASc Score

<table>
<thead>
<tr>
<th>CHADS2 Risk</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHF</td>
<td>1</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1</td>
</tr>
<tr>
<td>Age &gt; 75</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1</td>
</tr>
<tr>
<td>Stroke or TIA</td>
<td>2</td>
</tr>
</tbody>
</table>

### CHA₂DS₂-VASc Risk Score

<table>
<thead>
<tr>
<th>CHA₂DS₂-VASc Risk</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHF or LVEF ≤ 40%</td>
<td>1</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1</td>
</tr>
<tr>
<td>Age ≥ 75</td>
<td>2</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1</td>
</tr>
<tr>
<td>Stroke/TIA/Thromboembolism</td>
<td>2</td>
</tr>
<tr>
<td>Vascular Disease</td>
<td>1</td>
</tr>
<tr>
<td>Age 65 - 74</td>
<td>1</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
</tr>
</tbody>
</table>

From ESC AF Guidelines
### CHADS₂ vs. CHA²DS₂-VASc

<table>
<thead>
<tr>
<th>CHADS₂ score</th>
<th>Patients ($n = 1733$)</th>
<th>Adjusted stroke rate %/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>120</td>
<td>1.9</td>
</tr>
<tr>
<td>1</td>
<td>463</td>
<td>2.8</td>
</tr>
<tr>
<td>2</td>
<td>523</td>
<td>4.0</td>
</tr>
<tr>
<td>3</td>
<td>337</td>
<td>5.9</td>
</tr>
<tr>
<td>4</td>
<td>220</td>
<td>8.5</td>
</tr>
<tr>
<td>5</td>
<td>65</td>
<td>12.5</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>18.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHA²DS₂-VASc score</th>
<th>Patients ($n = 7329$)</th>
<th>Adjusted stroke rate %/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>422</td>
<td>1.3</td>
</tr>
<tr>
<td>2</td>
<td>1230</td>
<td>2.2</td>
</tr>
<tr>
<td>3</td>
<td>1730</td>
<td>3.2</td>
</tr>
<tr>
<td>4</td>
<td>1718</td>
<td>4.0</td>
</tr>
<tr>
<td>5</td>
<td>1159</td>
<td>6.7</td>
</tr>
<tr>
<td>6</td>
<td>679</td>
<td>9.8</td>
</tr>
<tr>
<td>7</td>
<td>294</td>
<td>9.6</td>
</tr>
<tr>
<td>8</td>
<td>82</td>
<td>6.7</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>15.2</td>
</tr>
</tbody>
</table>

Prognosis

- **Population:**
  - Recruitment strategy
  - Diagnosis confirmation,
  - Referral pattern,
  - Patient characteristics (demographic, comorbidities, vital signs for acute conditions, ...)
  - Inclusion & non-inclusion criteria
  - Treatment

- **Exposure/Comparison:**
  - variables, factors,
  - prognostic variables, predictors, risk factors, ..
  - Type of variables: clinical (CPR), biomarkers, imaging

- **Outcome:** events (single hard outcome, composite?) – drop-out

- **Time:** length of FU depends on condition
Prognosis

• **Study design:**
  – inception cohort,
  – prospective study
  – for prognosis

• **Sample size:** Events
  – Rule of thumb: 10 Events per candidate variable (10 EPV)
Prognosis research questions

• Proposal by PROGRESS group, ‘expert’ group on prognosis research

• Type I
  What is the basic variation in prognosis of this health related condition?

• Type II
  Is this prognostic factor association useful or causal?

• Type III
  How can we use multiple variables to make useful risk predictions in individuals?

• Type IV:
  Does this prognostic factor modify response to treatment?
Prognosis – Model development

- **Model construction** (multivariate analysis: regression models)
  - Derivation
  - Internal validation

- **Model validation** (external validation)
  - Geographic
  - Temporal

- **Model updating** (deleting variables, adding variables)

- **Model impact**
  - Management studies
  - RCT
Prognosis – Model Performance

• Distribution of risk-groups
• Incidence within risk-groups
  – Within low-risk group
  – Within high-risk group
• Classification (prognostic accuracy):
  – Se, Sp, LR-, LR+, PPV, NPV
• Reclassification
  – NRI
• Discrimination
  – ability to distinguish patients with different risks
  – ROC AUC
• Calibration
  – Does prediction agree with observation
  – How closely predicted data matches observed data
Prognosis – Critical appraisal

• Population (case ascertainment)
• Prognostic information (collection and ascertainment):
  – predictor variables (type, time, method of measurement)
  – FU (length sufficient enough, % lost to FU, reasons for lost to FU)?
  – Outcome (how defined – how assessed)
• Missing data (variables, patients, how treated)
• Confounders (in study design, in analysis, type)

Table 6.1  Is this evidence about prognosis valid?

1. Was a defined, representative sample of patients assembled at a common point in the course of their disease?
2. Was follow-up of study patients sufficiently long and complete?
3. Were objective outcome criteria applied in a “blind” fashion?
4. If subgroups with different prognoses are identified:
   • Was there adjustment for important prognostic factors?
   • Was there validation in an independent group of “test-set” patients?
Prognosis – Critical appraisal

<table>
<thead>
<tr>
<th>Table 6.2</th>
<th>Is this valid evidence about prognosis important?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>How likely are the outcomes over time?</td>
</tr>
<tr>
<td>2.</td>
<td>How precise are the prognostic estimates?</td>
</tr>
</tbody>
</table>

© Straus, Glasziou, Richardson, Haynes: Evidence-Based Medicine, 4th Edition.
Prognosis – Critical appraisal

Table 6.3  Can we apply this valid, important evidence about prognosis to our patient?

1. Is our patient so different from those in the study that its results cannot apply?
2. Will this evidence make a clinically important impact on our conclusions about what to offer or tell our patient?
References

References


Prognosis – Reporting Guidelines

EBM and Systematic Review

• **EBM**
  • Steps
    – Answerable Question
    – Search
    – Appraise
    – Apply
  • Time: 30 seconds

• **Systematic Review**
  • Steps
    – Answerable Question
    – Search ++++
    – Appraise x 2
    – Synthesize
    – Apply
  • Time: 6 months

Paul Glasziou
Systematic Review

http://www.cochrane.org/

• Research question
• Criteria for considering studies for the review
• Search methods for identification of studies
• Data collection and analysis
  – Selection of studies
  – Data extraction and management
  – Assessment of risk of bias
  – Measures of treatment effect
  – Unit of analysis issues
  – Dealing with missing data
  – Assessment of heterogeneity
  – Assessment of reporting bias
  – Data synthesis
  – Subgroup analysis and investigation for heterogeneity
  – Sensitivity analysis