INTRODUCTION TO TEST EVALUATION RESEARCH

Patrick MM Bossuyt

Overview

1. Tests
2. Evaluation of Tests – RCT
3. Target Condition
4. Test Accuracy Studies
5. Systematic Reviews of Test Accuracy
6. Coda

1. Tests

Why use tests?

- Diagnosis
- Monitoring course disease
- Selecting therapy
- Following effects of therapy
- Determining drug levels or drug effects
- Evaluate Health or Fitness
- Screening
- Case Finding
2. Evaluate Tests

Medical Test Information

Can I trust that this information tells me something about the patients' health?

Medical Test Information

Can I trust that this information improves patients' health?

Medical Test Information

Patient Outcome

Can I trust that this information helps improve patients' health?
**Randomized Clinical Trial**

Population → Study Group → Randomize → Active, Control → Outcome

**Medical Test RCT**

Population → Study Group → Randomize → TEST, Control → Outcome

**Medical Test RCT**

Population → Study Group → Randomize → No TEST, TEST B → Outcome

**Meta-Analysis Screening RCT**

<table>
<thead>
<tr>
<th>Task</th>
<th>No of studies</th>
<th>Meta Analysis (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>1, 2, 3, 4</td>
<td>0.84 (0.77 to 0.93)</td>
</tr>
</tbody>
</table>

**Combination antimicrobial susceptibility testing for acute exacerbations in chronic infection of Pseudomonas aeruginosa in cystic fibrosis**
Valerie Waters, Felix Ratjen
Year: 2008

**Screening for colorectal cancer using the faecal occult blood test, Hemoccult**
P Hewitson, P Glasziou, L Ireig, R Towler, E Watson
Year: 2007

**Urodynamic investigations for management of urinary incontinence in children and adults**
CMA Glazener, MC Lapitan
Year: 2002
Randomised controlled trial of faecal-occult-blood screening for colorectal cancer

Lancet 1996;348:1473

Summary

Background There is growing evidence that colorectal cancer screening programmes reduce colorectal cancer mortality by reducing cases to stage I and II cancer where curative local treatment is effective. In previous randomised trials, however, the effect of FOB screening on CRC mortality in such a setting has been null.

Methods Between February 1986 and January 1990, 174,949 people aged 50-64 years were invited to the study. Participants were randomly allocated FOB screening (29.4% were allocated to screening) to help in a community-based setting. The criteria were 1) the participant was invited for screening, and 2) the test was completed, and 3) the participant was followed up for cancer incidence and mortality. Those who refused to participate in the study were invited to be invited to the screening programme for CRC. The study was in a randomised group of 40,000 people from a general population. The follow-up time was 7.8 years.

Results

CRC incidence & mortality

<table>
<thead>
<tr>
<th>Rate (1000 years)</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening</td>
<td>1.49</td>
</tr>
<tr>
<td>Control</td>
<td>0.59</td>
</tr>
<tr>
<td>Total mortality</td>
<td>21.1</td>
</tr>
</tbody>
</table>

Median follow-up 7.8 years

Lancet 1996;348:1475

UK RCT:
CRC incidence & mortality

Randomised study of screening for colorectal cancer with faecal-occult-blood test

Lancet 1996;348:1473

Urodynamic investigations for management of urinary incontinence in children and adults

Lancet 1996;348:1468

Objective

The objective of this review was to discover if treatment according to a urodynamic-based diagnosis led to clinical improvements in urinary incontinence outcomes, compared to treatment based on history and examination.

Three small trials were found, which included 184 people, although information was only available for 128 participants. There was not enough evidence to determine whether these tests lead to better outcomes. There was some evidence that urodynamic testing increased the number of people prescribed drug treatments or treated by surgery, but it was not known whether this resulted in less incontinence or a better quality of life.

More research is needed, in which people are randomised to having treatment decisions based on either their symptoms and examination alone, or the extra information provided by urodynamic tests.
Tubal integrity testing

Intravenous Pyelography

Laparoscopic Procedure

www.womenshealthsection.com

Table 2: Characteristics of couples participating in study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intervention group (n=227)</th>
<th>Control group (n=217)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (y)</td>
<td>32.0 (30.6-34.4)</td>
<td>31.8 (30.5-33.5)</td>
</tr>
<tr>
<td>Mean duration of infertility (y)</td>
<td>5.9 (4.8-7.8)</td>
<td>6.0 (4.8-8.0)</td>
</tr>
<tr>
<td>Male infertility (y)</td>
<td>24.7 (21.9-27.5)</td>
<td>25.0 (22.3-27.7)</td>
</tr>
<tr>
<td>No (%) couples with history of renal problems</td>
<td>10 (5)</td>
<td>12 (6)</td>
</tr>
<tr>
<td>No (%) couples with bowel problems</td>
<td>20 (9)</td>
<td>21 (9)</td>
</tr>
<tr>
<td>No (%) of women with uterine disorder</td>
<td>34 (15)</td>
<td>36 (17)</td>
</tr>
<tr>
<td>No (%) of women with pelvic disorder</td>
<td>30 (13)</td>
<td>37 (17)</td>
</tr>
</tbody>
</table>

Cumulative pregnancy rates for 227 couples in intervention group (which included postcoital test) and 217 couples in control group (which excluded the test)
Field interventional Radiology: randomizing participants in a randomized controlled trial (RCT) for medical tests.

**RCT Medical Test**

- **Study Group Randomize**
- **Test**
- **No Test**
- **Outcome**
- **Test A Result**
  - **Left Positive**
  - **Right Negative**
  - **Outcome**
  - **Test B Result**
  - **Left Positive**
  - **Right Negative**
  - **Outcome**

**DSA versus Multi-Detector Row CT Angiography in Peripheral Arterial Disease: Randomized Controlled Trial**

- **Randomized**
- **111 eligible patient with symptomatic PAOD**
  - **52 not randomized**
    - 50 patient refusals
    - 2 patients moved
  - 15 emergency cases
  - 2 technical failures

- **75 MDCTA group**
  - 45 MDSA group
  - 1000 ultrasound
  - 1 time

- 11 at 12-month follow-up assessment: symptomatic improvement
  - need for additional imaging and/or intervention

RCT of Testing

- Need patient outcomes that matter
- Need protocol
- Usually need large sample sizes
- RCT of Testing
- Best evidence of effectiveness
Spectrum of disease

Venous Thromboembolism

Roles of tests

4. Test Accuracy
Bladder Tumor Markers (BTM)

- To optimize monitoring of tumor recurrence and progression, without incurring more invasive and expensive medical tests.

Diagnostic Accuracy

- Index Test
- Reference Standard

Diagnostic Accuracy Study

- Series of Patients
- Index Test
- Reference standard
- Cross-classification

The Results

<table>
<thead>
<tr>
<th>Reference Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index Test</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Negative</td>
</tr>
</tbody>
</table>

Bladder Cancer

- Close follow-up due to the significant risk of tumor recurrence.

Cytology: central test

- Efficient
- Non-invasive
- Inexpensive
- But imperfect
- Cystoscopy
To optimize monitoring of tumor recurrence and progression, without incurring more invasive and expensive medical tests.

Example

Diagnostic Accuracy Study

- 501 patients
- BTA stat
- Cystoscopy
- Cross-classification

The Results

<table>
<thead>
<tr>
<th>Cystoscopy</th>
<th>BCR</th>
<th>No BCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>71</td>
<td>96</td>
</tr>
<tr>
<td>Negative</td>
<td>62</td>
<td>272</td>
</tr>
</tbody>
</table>

Measures of Diagnostic Test Accuracy

- Sensitivity & Specificity: 54% & 74%
- PPV & NPV: 43% & 81%
- Likelihood Ratios: 2.0 & 0.63
- Diagnostic Odds Ratio: 3.2

<table>
<thead>
<tr>
<th>BCR</th>
<th>No BCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>71</td>
</tr>
<tr>
<td>Negative</td>
<td>62</td>
</tr>
</tbody>
</table>
Message

- Measures of Diagnostic Test Accuracy are test characteristics, not fixed test properties.

5. Systematic Reviews

Why systematic reviews

- Extensive/exhaustive search
- Critical appraisal
- Meta-analysis
  - Increased precision
  - Explore variability

1. Focused question

Systematic Review: Question Elements

What is the diagnostic accuracy of

<index test>

[versus <comparator>]

for detecting <target condition>

in <patient description>

Systematic Review: Question Elements

What is the diagnostic accuracy of

BTA stat test

versus cytology

for detecting recurrent disease

in patients with bladder cancer
ROC space

2. Identification & selection studies

3. Quality Assessment

Differential Verification Bias

4. Meta-analysis
Study results

Paired Forest Plots

ROC space
Summary Point in ROC space

Summary Point in ROC space

Summary Point in ROC space

Summary Point in ROC space

Summary Curve in ROC space

Paired studies: cytology & BTA
5. Interpretation