Going from evidence syntheses to recommendations about diagnostic questions: When can test accuracy results be sufficient to extrapolate about overall benefits and harms

Reem Mustafa, MD, MPH, PhD
June 2016
Study designs for diagnosis

• If a test fails to improve important outcomes (incl. reducing resources, complications): no reason to use it, whatever its accuracy

• Recognized for over 20 years

• Best way to assess diagnostic strategy: randomized controlled trial or well-done observational studies in which investigators focus on patient-important outcomes

• However, poor guidance for what to do if no such evidence available
Are there studies that directly focus on: mortality, morbidity, symptoms, and/or quality of life?

Apply GRADE approach as for treatment or other intervention

Schunemann et al. BMJ, 2008
Study designs II

Look for diagnostic test accuracy studies

And then draw inferences from other evidence
Accuracy

sensitivity & specificity, LR+ & LR-

Patient-important consequences

TP  (treated...)
TN  (reassured...)
FP  (needlessly treated...)
FN  (not treated...)

Inconclusive results

Complications
Test accuracy is a surrogate for patient important outcomes

• When clinicians think about diagnostic tests, they focus on their accuracy

• Underlying assumption: obtaining a better idea of whether a target condition is present or absent will result in superior patient management and improved outcome.
longer life
fewer symptoms
fewer complications
better quality of life

sensitivity
specificity

surrogate
So what should we do when making recommendations about diagnostic tests when RCT or observational studies about tests’ effect on patient important outcomes are lacking?
Apply GRADE for Diagnostic Test Accuracy studies

5 factors to downgrade (factors to upgrade?)

- High
- Moderate
- Low
- Very low

Summary of findings & estimate of effect for each outcome

Create DTA evidence synthesis (pooled estimates of DTA)
Should HPV or VIA be used to screen for CIN 2+ in asymptomatic women in LMIC?

**Asymptomatic women**

- **HPV**
  - Test + (TP & FP)
    - Cryo eligible?
      - No: Treat with LEEP
      - Yes: Treat with Cryo
      - Outcomes*
  - Test - (TN & FN)

- **VIA**
  - Test + (TP & FP)
    - Cryo eligible?
      - No: Treat with CKC
      - Yes: Treat with Cryo
      - Outcomes*
  - Test - (TN & FN)

*Outcomes*
Apply GRADE for Diagnostic Test Accuracy studies

5 factors to downgrade (factors to upgrade?)

- High
- Moderate
- Low
- Very low

DIRECTNESS OF THE OUTCOME

Lower our confidence

Very uncertain
Uncertain
Moderately certain
Certain
Very certain

Diagnos2c Test Accuracy

Quality of evidence

Sensitivity
Specificity

STUDIES

NATURAL HISTORY

Patients will suffer from disease without being detected or suffer from symptoms and undergo repeat testing or testing for other disease that will happen at certain rate.

SR required, full framework always needs to be developed.

Mortality
Morbidity
QoL
Harms
Resources
Other

Final Quality of evidence for each outcome based on DTA and linked evidence and development of recommendations

High
Moderate
Low
Very low

e.g. 3%
Apply GRADE for Diagnostic Test Accuracy studies

5 factors to downgrade (factors to upgrade?)

High ⊕⊕⊕⊕
Moderate ⊕⊕⊕O
Low ⊕⊕OO
Very low ⊕OOU

Studies assessing management outcomes in TP who will, e.g., be treated. All relevant therapeutic strategies would need to be considered, effects evaluated and used for modelling.

Directness of the outcome

Mortality
High ⊕⊕⊕⊕
Moderate ⊕⊕⊕O
Low ⊕⊕OO
Very low ⊕OOU

Morbidity

QoL

Harms

Resources

Other

Assess linked evidence

Final Quality of evidence for each outcome based on DTA and linked evidence and development of recommendations

Diagnos2c Test Accuracy Synthesis

Sensitivity

Specificity

TP
FP
TN
FN
Can test accuracy results be sufficient to extrapolate about overall benefits and harms?

- Literature survey
  - DTA studies
  - DTA SR
  - Guidelines about tests

- Experts interviews
**Test accuracy results may be sufficient to extrapolate about overall benefits and harms**

<table>
<thead>
<tr>
<th>“Win-win” situation when all else is equal or better.*</th>
<th>Replacing culture by PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic non-inferiority is sufficient for a decision when inferences can be made about the impact on patient important outcomes</td>
<td>New” HIV diagnostic test is non-inferior to an existing HIV diagnostic test, given that benefits of treatment of HIV are already established beyond doubt</td>
</tr>
<tr>
<td>When the test accuracy of one test is equivalent or better than the combined accuracy of two tests (one of which is the test evaluated separately)</td>
<td>Plain radiography does not add any clinically significant advantage to multidetector row CT</td>
</tr>
<tr>
<td>When the primary goal is to establish a diagnosis for a condition or rule out a diagnosis **</td>
<td>Some people want to look at their babies before they are born with ultrasound even if they have no intention of doing anything about abnormalities they might detect</td>
</tr>
</tbody>
</table>