Institute of Applied Health Sciences

Dr Sue Mallett (s.mallett@bham.ac.uk)
21st July 2016
Measuring the impact of diagnostic tests on patient management decisions within three clinical trials

- Sue Mallett (University of Birmingham)
- Stuart Taylor (CI of trials, UCL)
- Gauraang Batnagar (MD on METRIC trial, UCL)
- STREAMLINE COLON Investigators
- STREAMLINE LUNG Investigators
- METRIC Investigators
Overview of presentation

• Background to diagnostic impact studies
• Difficulties with RCT designs
• Alternative designs
• Background to three trials measuring patient management
• Approaches in these three trials
• Issues with biases
• Future studies and research
How do diagnostic tests impact on patient outcomes?

Criticism of diagnostic accuracy studies
- Want to know more than diagnostic accuracy
- Want impact of using a test on patient management and patient outcomes

How do we study impact of test on patient outcomes?

Test → Patient outcomes

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Management decision</th>
<th>Treatment efficacy</th>
</tr>
</thead>
</table>

RCT design: Test as an intervention

Best option: when feasible

| Diagnostic accuracy | Management decision | Treatment efficacy |

All three components are combined as complex intervention
RCT design for diagnostic test as invention?

Problems with RCT
If no effect of test on patient outcomes reason is unclear
  • Test is inaccurate?
  • Patient management was inconsistently based on test?
  • Treatment is not effective?
Better test will have no impact if treatment not effective… but if next year there may be an effective treatment
  • We do not know if diagnostic test could be useful
Other difficulties in RCT designs

- If clinicians have no confidence in test, then patient management will not change (and shouldn’t!)
- Outcome studies take long time e.g. for cancer may take 5 or 10 years of follow up
  - Diagnostic tests, patient pathways and treatments change rapidly
  - RCT could be a trial of what you should have done 5 or 10 years ago, when now there is a different test or different clinical pathway
- Sample sizes for RCTs are often huge
- Some trials are not ethical
- Some trials are impossible to recruit
Diagnosis: New approach in evaluating impact of tests

Best option: when feasible

RCT of test as intervention (when feasible)

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Management decision</th>
<th>RCT of treatment</th>
</tr>
</thead>
</table>

Other options: maybe best can do

Usual method

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Modelling based on little or no data</th>
<th>RCT of treatment</th>
</tr>
</thead>
</table>

Diagnostic impact on patient management

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Management decision</th>
<th>Modelling</th>
<th>RCT of treatment</th>
</tr>
</thead>
</table>

UNIVERSITY OF BIRMINGHAM
Measuring management decisions

Diagnostic accuracy → Management decision → Modelling → RCT of treatment

Diagnostic impact on patient management

Advantages

• Faster – so study faster and cheaper, and test might even be the same at the end of the trial…
• Can look at component parts e.g. test accuracy
• Can update for new treatment

Disadvantages

• New type of design with additional biases
• Will only be realistic where clinicians confident to use test
• Need to include treatment from modelling studies
Three example trials

- **METRIC**: comparing US and MRI in diagnosis (staging) of Crohn’s disease
- **STREAMLINE COLON and LUNG**: comparing whole body MRI to standard NICE pathways (CT etc) for detecting metastases in colorectal and lung cancer
METRIC: Crohn's disease – location and activity of disease

Patients with new diagnosis or relapse of disease

- Chronic inflammatory bowel disease diagnosed < 25 yrs
- Requires lifelong medical and surgical therapy
- Affects 150,000 people in the UK (around 1 in 700)
- 27,000 hospital admissions annually (Leiper K 2006)
- CT is usual imaging test for diagnosis but has radiation exposure and cancer risk
- Tests without radiation risk would be preferred (US, MRI)

Ultrasound

MRI

Reference standard

Accuracy of diagnosis

Impact of test on patient management
STREAMLINE colon: detection of metastasis

- Patients with new diagnosis of colorectal cancer
- Test of interest: Early whole body MRI
- Comparison: NICE pathway (CT plus MRI and PET as needed)
- Reference standard: 12m FU all tests
- Accuracy of diagnosis
- Impact of test on patient management
  - Reported in real time in patient pathway by multidisciplinary team meeting (MDT)
STREAMLINE: How to capture patient management?

Within real clinical patient pathway

Accurate:

Radiologists blinded

High applicability

High potential for bias: lack of blinding

Modeling to RCT of treatment

Index test 1

Patient management decision 1

Index test 2

Patient management decision 2

UNIVERSITY OF BIRMINGHAM
METRIC: How to capture patient management?

Theoretical patient management decisions (outside of real patient pathway)

- Index test 1
- Index test 2
- Patient management decision 1
- Patient management decision 2

Accuracy:
- Radiologists blinded
- Low bias but low applicability: not real clinical pathway?

Modelling to RCT of treatment
Issues in design

• To blind between patient management decisions from compared index tests
  - Requires different clinicians making management decisions or separation by time (so not real patient pathway)
• On or off real patient pathway?
  - Same urgency
  - Same type of clinicians
• What information will be included in patient management decision?
  - affected by on or off patient pathway?
• What to do about diseases where big differences in patient management between clinicians?
Applicability and risk of bias

Often risk of bias and applicability pull in opposite directions.
Other issues in design

• Will clinicians use test to direct patient management?
  ➢ Needs to be established test

• Reference standard for management?
  ➢ OK for STREAMLINE colon and lung, as cancer treatments are internationally agreed.
  ➢ Problem in Crohn’s disease as treatment varies between clinicians

• How to analyse? Percentage change or rank of seriousness of change?
Diagnosis: New approach in evaluating impact of tests

**Best option: when feasible**

RCT of test as intervention (when feasible)

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Management decision</th>
<th>RCT of treatment</th>
</tr>
</thead>
</table>

**Other options: often best can do**

**Usual method**

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Modelling based on little or no data</th>
<th>RCT of treatment</th>
</tr>
</thead>
</table>

**Diagnostic impact on patient management**

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>Management decision</th>
<th>Modelling</th>
<th>RCT of treatment</th>
</tr>
</thead>
</table>
Current and future research

- Recording patient management a bonus add-on to commissioned diagnostic accuracy studies
- Fourth study just starting as add-on diagnostic study…
- Three variations in study design in 4 trials
- Future HTA commissioned research likely to include management
- Methodology work included in and alongside these studies (unfunded despite applications for funding)
Thank you and questions
Blinding patient management decisions

- Ethically all test results need to be used to make treatment decisions for patients

Options for blinding

1. **Within patient clinical pathway (STREAMLINE, MROC)**
   - Multiple MDTs blinded within patient clinical pathway? Could delay and mess up patient care
   - Randomise to different tests revealed first. NICE or WB-MRI? Complicates and doubles sample size
   - Accept lack of blinding – Instruct MDT to use radiologist reports (which are blinded and trial statistician can check against).

2. **Outside of patient clinical pathway – theoretical decisions (METRIC, MROC)**
   - Randomise to different tests revealed. >1 month gap between each test assessment
Some important biases in diagnostic studies

- **Blinding of test interpretation** – are interpretations blinded (i) index to reference (ii) reference to index?
- **Experience of those reading test** – Typical of clinical practice?
- **Clinical spectrum of patients** – are patients typical of those receiving test in clinical practice?
- **Verification bias** – does everyone get a reference test or same reference test?
- **Incorporation bias** – does index test form part of reference test?
- **Time between tests** – Are tests within time period where condition does not change?
Sample size for RCT tests as interventions: why so large?

- Effective sample size based on number of patients who would have a different diagnosis (10%?).
- But not all of these may get a different patient management (half of patients?)
- Could be 20 times larger sample size or more than an RCT of treatment?
### Blinding: Diagnostic test result

<table>
<thead>
<tr>
<th>Index test accuracy</th>
<th>STREAMLINE</th>
<th>METRIC</th>
<th>MROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDT (but on basis of blinded radiologist interpretation) then MRI. Must be MDT, as multi-test pathway*</td>
<td>blinded radiologist interpretation</td>
<td>CT online in patient pathway (with and without critical findings from MRI) MRI off line (blinded to CT)**</td>
<td></td>
</tr>
</tbody>
</table>

* Have individual test results, so can understand what component of final decision.

** Methodological interest: Is this the same as CT with critical findings from MRI?
<table>
<thead>
<tr>
<th>Patient management</th>
<th>STREAMLINE</th>
<th>METRIC</th>
<th>MROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDT. Fixed order, MRI not blinded.</td>
<td>Blinded offline interpretation, &gt;1 month gap, random order</td>
<td>Blinded offline interpretation, &gt;1 month gap, random order</td>
<td></td>
</tr>
</tbody>
</table>

UNIVERSITY OF BIRMINGHAM
Issues as trial is ongoing

Unblinding of capability of index tests as trial is ongoing

• Consensus panels (reference standard) necessarily use same radiologists

• Clinical team treating patients – unethical to withhold information that could benefit patients

Note accumulation of data across patients is kept unblinded