Teaching Tips for Diagnostic Studies

Dr. Annette Plüddemann

Department of Primary Care Health Sciences
Centre for Evidence-Based Medicine
Horizon Scanning

NON-CONTACT INFRARED THERMOMETERS 0025

Download report here
Read full report here

POINT-OF-CARE (POC) TESTING FOR A PANEL OF CARDIAC MARKERS 0033

Download report here
Read full report here

POINT-OF-CARE CALPROTECTIN TESTS 0034

Download report here
Read full report here

POINT-OF-CARE FAecal OCCULT BLOOD TESTING 0035

Download report here

The Oxford Diagnostic Horizon Scan Programme identifies new and emerging diagnostic technologies relevant to primary care in the NHS.

Aims:
1. to identify new and emerging diagnostic technologies relevant to primary care in the NHS that are likely to have significant impact
2. to filter and prioritise these diagnostic technologies
3. to summarise the evidence base for technologies
4. to make an early assessment of their likely impact on provision of health care and health care outcomes
5. to highlight and outline research requirements to facilitate implementation.

We use a variety of methods to search for new technologies, such as literature searches and interaction with the diagnostics industry and clinicians, and prioritise certain technologies to produce technology reports, systematic reviews and health economic assessments. In some cases we conduct pilot studies of new technologies in primary care settings. The technologies we identify may be relevant to diagnosis as well as monitoring, assessing...
Search results for: EBHC
Currently displaying courses 1 to 10 of 10

Subject categories
- All subjects (10)
  - Health (10)

Search results:

1. **The History and Philosophy of Evidence-Based Health Care**
   - Course code: CHN01X
   - Type: Professional Development course
   - Duration: 15 Jun 2015
   - Description: An exploration of the history and philosophy of Evidence-Based Health Care, including 'placebo controls', 'the evolution of scientifically-defensible methods of information synthesis' and 'whether average results apply to individuals'.

2. **Evidence-Based Diagnosis and Screening**
   - Course code: CHN02X
   - Type: Professional Development course
   - Duration: 26 Jan 2015
   - Description: The aim of this module is to critically appraise and apply, at an advanced level, the best evidence on diagnostic tests. This will briefly revise then extend material on diagnostics from the introductory modules.

3. **MSc in Evidence-Based Health Care**
   - Course code: CHN03X
   - Type: Professional Development course
   - Duration: 13 Oct 2014
   - Description: The MSc in Evidence-Based Health Care is part of the Oxford International Programme in Evidence-Based Health Care, and is offered as a part-time course consisting of six taught modules and a dissertation.

4. **Mixed Methods in Health Research**
   - Course code: CHN04X
   - Type: Professional Development course
   - Duration: 23 Feb 2015
   - Description: This module will introduce and investigate the nature and design of mixed methods research in evidence-based health care.

5. **Practice of Evidence-Based Health Care**
   - Course code: CHN05X
   - Type: Professional Development course
   - Duration: 16 Mar 2015
   - Description: This module will introduce the basic concepts and skills of Evidence-Based Health Care. Teaching will be led by Dr Carl Henneghan and Dr Annette Flüddemann.

6. **The Practice of Evidence-Based Health Care (Surgery)**
   - Course code: CHN06X
   - Type: Professional Development course
   - Duration: 13 Oct 2014
   - Description: This course introduces the basics of scientific methodology and study design as applied to clinical research. Teaching takes place over five days and is led by Mr Peter McGuire, Dr Carl Henneghan and Dr Annette Flüddemann.
So far….

Randomised controlled trial

Systematic review of an intervention

Is this treatment better than doing nothing?

Which treatment is better?
Typically someone with abnormal *symptoms* consults a physician, who will obtain a history of their illness and examine them for *signs* of diseases.

The physician formulates a hypothesis of likely diagnoses and may or may not order further *tests* to clarify the diagnosis.
Diagnostic Errors—The Next Frontier for Patient Safety

David E. Newman-Toker, MD, PhD
Peter F. Pronovost, MD, PhD

During the past decade, awareness and understanding of medical errors have expanded rapidly, with an energetic patient safety movement promoting safer health care through “systems” solutions. Efforts have focused on translating evidence into practice, mitigating hazards from therapies, and improving culture and communication. Diagnostic errors have received relatively little attention. Although the science of error measurement is underdeveloped, diagnostic errors are an important source of preventable harm. In this Commentary, we offer definitions for diagnostic error and misdiagnosis-related harm, present an overview of the magnitude of diagnostic errors, and give suggestions for how research can mature.

Distinguishing Errors From Harms

In considering diagnostic errors, it is important to distinguish between the error (a process) and the resulting harm (an outcome). Diagnostic error can be defined as a diagnosis that is missed, wrong, or delayed, as detected by some subsequent definitive test or finding. However, all misdiagnoses result in harm, and harm may be due to either disease or intervention. Misdiagnosis-related harm can be defined as preventable harm that results from the delay in failure to treat a condition actually present (when the workup diagnosis is wrong or unknown), or from treatment provided for a condition not actually present.

An estimated 40,000 to 80,000 US hospital deaths result from misdiagnosis annually. Roughly 3% of autopsies reveal lethal diagnostic errors for which a correct diagnosis coupled with treatment could have averted death. In the Harvard Medical Practice Study, physician errors resulting in adverse events were more likely to be diagnostic than drug-related (14% vs 9%), and misdiagnoses were more likely to be considered negligent (73% vs 53%) and to result in serious disability (47% vs 14%). Not surprisingly, tort claims for diagnostic errors are nearly twice as common as claims for medication errors and result in the largest payouts. With all types of medical error, the human toll of misdiagnoses on an individual or family can be tremendous, particularly when a healthy patient experiences an adverse event. Diagnostic errors often are unrecognized or unreported, and the science of measuring these errors (and their effects) is underdeveloped.

Available statistics consider neither deaths due to misdiagnoses in outpatients nor misdiagnosis-related morbidity and associated costs. For example, stroke, the leading cause of serious, long-term disability in the United States, afflicts 780,000 Americans annually. Opportunities to prevent disabling stroke are missed when patients experiencing mild or transient warning symptoms receive misdiagnoses. According to a recent systematic review, 99% of all cerebrovascular events are missed initially, and the odds of misdiagnosis increase as time lapse beyond symptoms are mild or transitory.

Author Affiliations: Departments of Neurology (D Newman-Toker and AM Hauer) and Obstetrics and Gynecology (P F Pronovost), Johns Hopkins University School of Medicine, Baltimore, Maryland.

Corresponding Author: David E. Newman-Toker, MD, PhD, The Johns Hopkins Hospital, Department of Neurology, 600 North Wolfe Street, Baltimore, MD 21287 (tokerdj@jhmi.edu).

The Clinical Evidence Bulletin (CEBM) is the online health information database. It provides critically evaluated evidence to support health care decisions. CEBM is a joint product of the University of Oxford, the University of Utah, and the Mayo Clinic. It is supported by the National Library of Medicine and the National Institutes of Health. CEBM is an independent, non-profit organization. CEBM is available at www.cebm.net. (03/2008 American Medical Association. All rights reserved.)

- 2/3 legal claims against GPs in UK
- 40,000-80,000 US hospital deaths from misdiagnosis per year
- Adverse events, negligence cases, serious disability more likely to be related to misdiagnosis than drug errors
- Diagnosis uses <5% of hospital costs, but influences 60% of decision making
Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study

Helen Hogan,1 Frances Healey,2 Graham Neale,3 Richard Thomason,4 Charles Vimpani,5 Nick Black1

ABSTRACT: Monitoring hospital mortality rates is widely recommended. However, the number of preventable deaths remains uncertain with estimates in England ranging from 8,400 to 48,000 per year. These being derived from studies that identified adverse events but not whether events contributed to death or shortened life expectancy of those affected.

Methods: Retrospective case record reviews of 1000 adults who died in 2009 in 16 acute hospitals in England were undertaken. Trained physician reviewers independently identified 255,000 NHS patients each year suffer serious disability or death as a result of healthcare interventions. This estimate was derived from retrospective case record reviews (RCCRs) studies conducted in USA in the 1980s and 90s.4 These and other national studies using comparable methods were not designed to establish the proportion of deaths that were preventable.5-8 Two smaller studies have specifically assessed the degree to which problems in

• clinical monitoring (such as failure to act upon test results or monitor patients appropriately) – identified as a problem in 31% of preventable deaths
• diagnosis (such as problems with physical examination or failure to seek a specialist opinion) – identified as a problem in 30% of preventable deaths
• drugs or fluid management – identified as a problem in 21% of preventable deaths
Pensioner told she had dementia and sold her house to pay for care is told 18 months later there is NOTHING wrong with her

- Winnie Hill, 88, was distraught after doctors told her she had Alzheimer's disease
- Family were advised she would need round-the-clock care in special home
- Mrs Hill's daughter sold her house to fund the new accommodation
- She hated living in new home and said she didn't believe she had dementia
- After 18 months, her worried daughter sought the opinion of another doctor
- He declared she had mild cognitive impairment - not the same as dementia
- 'Furious' family moved her out of specialist home, and Mrs Hill now happier

By ANNA HODGEKISS FOR MAIL ONLINE
Phone apps may delay skin cancer diagnosis

Using a smartphone app to decide whether a mole is cancerous could delay sometimes life-saving treatment, according to American researchers.

The University of Pittsburgh scientists put four applications to the test by showing them 186 pictures of cancers and less concerning skin conditions.

Three of the apps wrongly labelled the cancerous lesions as unproblematic in almost a third of cases.

Doctors warn using phones rather than seeking expert help could be harmful.

The research, published in the journal JAMA Dermatology, looked at four commonly used applications.

The images selected to test the apps were all of skin lesions that were later removed and checked for an accurate diagnosis.

Three of the apps analysed the pictures using automated algorithms, without the involvement of doctors.

But users submitting pictures to the fourth app had their images reviewed by a qualified skin specialist.

In this case only one out of 53 cancerous lesions was misdiagnosed, but this app cost $5 (£3.10) per use.

Prof Laura Ferris, lead researcher of the study, said: "It is important that patients consult a doctor if they are concerned about a mole."

Related Stories

Drug holidays' heat cancer defence

'I got skin cancer in my 20s'

Skin cancer: Mole warning signs

Most Popular

The myth of working from home

The false alarm
announced by clinics to 18 months

• More than 400,000
• Estimates foreca.
• Rate of successful

By TIM SHIPMAN
PUBLISHED: 22:48, 4 November 2012

Comments (137) | Share

Political drive to screen for pre-dementia: not evidence based and ignores the harms of diagnosis

David G La Couteur professor of geriatric medicine 1, Jenny Doust professor of clinical epidemiology 2, Helen Creasey dementia specialist 3, Carol Brayne professor of public health 4

Centre for Education and Research on Ageing, ANZAC Medical Research Institute and the Charles Perkins Centre, University of Sydney and Sydney Research, Concord, 2139, Australia; Centre for Research in Evidence Based Practice, Bond University, Robina, Australia; Centre for Education and Research on Ageing, Concord RHO Hospital; Department of Public Health and Primary Care, University of Cambridge, Cambridge,

Current policy in many countries is aimed at increasing the rate of diagnosis of dementia and cognitive impairment. This policy has been accompanied by research into early detection of dementia, including preclinical identification of underlying neurobiology that might later be associated with dementia. Evidence to support their use. Little attention has been paid to the fact that attending memory clinics generates stress for patients and their carers, and expands the use of biomarker testing (cerebrospinal fluid measurements of amyloid and tau) and neuroimaging, which has associated costs and morbidity.
# Module Timetable

<table>
<thead>
<tr>
<th>Time</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.30</td>
<td>Registration</td>
<td>Registration</td>
<td>Registration</td>
<td>Registration</td>
<td>Registration</td>
</tr>
<tr>
<td>9.00-10.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.30-11.00</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
</tr>
<tr>
<td>11.00-12.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.30-13.30</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
</tr>
<tr>
<td>13.30-15.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.00-15.30</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
</tr>
<tr>
<td>15.30-17.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Monday 27 January</td>
<td>Tuesday 28 January</td>
<td>Wednesday 29 January</td>
<td>Thursday 30 January</td>
<td>Friday 31 January</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>8.30</td>
<td>Registration</td>
<td>Registration</td>
<td>Registration</td>
<td>Registration</td>
<td>Registration</td>
</tr>
<tr>
<td>9.00-10.30</td>
<td>Introduction Carl Heneghan&lt;br&gt;Annette Plüddemann&lt;br&gt;Diagnostic studies — the numbers Annette Plüddemann&lt;br&gt;Visualising and presenting diagnostic accuracy studies Susan Mallett&lt;br&gt;Multiple tests Ann Van den Bruel&lt;br&gt;Rational monitoring Jason Oke and Richard Stevens (9.30 start)</td>
<td>Diagnostic studies — the numbers Annette Plüddemann&lt;br&gt;Visualising and presenting diagnostic accuracy studies Susan Mallett&lt;br&gt;Multiple tests Ann Van den Bruel&lt;br&gt;Rational monitoring Jason Oke and Richard Stevens (9.30 start)</td>
<td>Diagnostic study design Ann Van den Bruel&lt;br&gt;Sample size calculations Richard Stevens&lt;br&gt;Systematic reviews of diagnostic studies Clare Bankhead&lt;br&gt;Screening Paul Hewitson&lt;br&gt;Rational monitoring Jason Oke and Richard Stevens (9.30 start)</td>
<td>Rational monitoring Jason Oke and Richard Stevens (9.30 start)</td>
<td>Rational monitoring Jason Oke and Richard Stevens (9.30 start)</td>
</tr>
<tr>
<td>10.30-11.00</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
</tr>
<tr>
<td>11.00-12.30</td>
<td>Diagnostic study design Ann Van den Bruel</td>
<td>Sample size calculations Richard Stevens</td>
<td>Systematic reviews of diagnostic studies Clare Bankhead</td>
<td>Screening Paul Hewitson</td>
<td>Rational monitoring Jason Oke and Richard Stevens (9.30 start)</td>
</tr>
<tr>
<td>12.30-13.30</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
</tr>
<tr>
<td>13.30-15.00</td>
<td>Searching skills Nia Roberts&lt;br&gt;Appraisal workshop Ann Van den Bruel</td>
<td>Systematic reviews of diagnostic studies (2) Clare Bankhead</td>
<td>Innovation and Adoption of Diagnostic Services Chris Price</td>
<td>Student presentations</td>
<td>Student presentations</td>
</tr>
<tr>
<td>15.00-15.30</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
<td>Coffee</td>
</tr>
<tr>
<td>15.30-17.00</td>
<td>Own search</td>
<td>Intermediate, indeterminate and uninterpretable results Beth Shinkins&lt;br&gt;Meta-analysis Beth Shinkins and Ann Van den Bruel&lt;br&gt;Consolidation and Q&amp;A session Carl Heneghan&lt;br&gt;Evaluation and finish by 4.30</td>
<td>Meta-analysis Beth Shinkins and Ann Van den Bruel&lt;br&gt;Consolidation and Q&amp;A session Carl Heneghan&lt;br&gt;Evaluation and finish by 4.30</td>
<td>Meta-analysis Beth Shinkins and Ann Van den Bruel&lt;br&gt;Consolidation and Q&amp;A session Carl Heneghan&lt;br&gt;Evaluation and finish by 4.30</td>
<td>Meta-analysis Beth Shinkins and Ann Van den Bruel&lt;br&gt;Consolidation and Q&amp;A session Carl Heneghan&lt;br&gt;Evaluation and finish by 4.30</td>
</tr>
<tr>
<td>18.00-19.00</td>
<td>Invited guest speaker: Beth Shinkins</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.00</td>
<td>Drinks reception, Rewley House</td>
<td></td>
<td></td>
<td></td>
<td>Dinner: Rewley House</td>
</tr>
</tbody>
</table>
Diagnostic strategies and what tests are used for
How do clinicians make diagnoses?

• Patient history…examination…differential diagnosis…final diagnosis

• Aim: identify types and frequency of diagnostic strategies used in primary care
  – 6 GPs collected and recorded strategies used on 300 patients.

Diagnostic stages & strategies

Stage

Initiation of the diagnosis

Refinement of the diagnostic causes

Defining the final diagnosis

Strategies used

• Spot diagnoses
• Self-labelling
• Presenting complaint
• Pattern recognition

• Restricted Rule Outs
• Stepwise refinement
• Probabilistic reasoning
• Pattern recognition fit
• Clinical Prediction Rule

• Known Diagnosis
• Further tests ordered
• Test of treatment
• Test of time
• No label

(Heneghan et al, BMJ 2009)
Not all diagnoses need tests?

Spot diagnosis

Meningitis

Chicken Pox
What are tests used for?

- Increase certainty about presence/absence of disease
- Disease severity
- Monitor clinical course
- Assess prognosis – risk/stage within diagnosis
- Plan treatment e.g., location
- Stall for time!

“Off hand, I'd say you're suffering from an arrow through your head, but just to play it safe, I'm ordering a bunch of tests.”
Roles of new tests

- **Replacement** – new replaces old
  - E.g. CT colonography for barium enema
- **Triage** – new determines need for old
  - E.g. B-natriuretic peptide for echocardiography
- **Add-on** – new combined with old
  - E.g. ECG and myocardial perfusion scan

Bossuyt et al BMJ 2006;332:1089–92
Critical appraisal of a diagnostic accuracy study
Diagnostic tests: What you need to know

• Validity of a diagnostic study

• Interpret the results
Defining the clinical question: PICO or PIRT

- **Patient/Problem**
  How would I describe a group of patients similar to mine?

- **Index test**
  Which test am I considering?

- **Comparator… or …Reference Standard**
  What is the best reference standard to diagnose the target condition?

- **Outcome….or….Target condition**
  Which condition do I want to rule in or rule out?
Diagnostic Accuracy Studies

Series of patients

Index test

Reference standard

Compare the results of the index test with the reference standard, blinded
Near patient testing for influenza in children in primary care: comparison with laboratory test

Anthony Harnden, Angela Brueggemann, Sasha Shepperd, Judy White, Andrew C Hayward, Maria Zambon, Derrick Crook, David Mant

Influenza is an important cause of acute respiratory illness in young children. Common complications include febrile convulsions, otitis media, bronchiolitis, and croup. In epidemic years attack rates among preschool children often exceed 40%. During these years children with influenza may account for up to 30% of the increase in antibiotic prescribing. Symptoms and signs of influenza in children are not specific and can mimic a range of other common respiratory viral pathogens. One quick way of reaching a precise diagnosis in primary care is to use a near

<table>
<thead>
<tr>
<th></th>
<th>RT-PCR test</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Near patient test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>27</td>
<td>3</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>94</td>
<td>93</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>96</td>
<td>157</td>
<td></td>
</tr>
</tbody>
</table>
Appraising diagnostic studies: 3 easy steps

Are the results valid?
• Appropriate spectrum of patients?
• Does everyone get the reference standard?
• Is there an independent, blind or objective comparison with the reference standard?

What are the results?

Will they help me look after my patients?
1. *Appropriate spectrum* of patients?

- Ideally, test should be performed on a group of patients in whom it will be applied in the real world clinical setting.

- **Spectrum bias** = study using only highly selected patients......perhaps those in whom you would really suspect have the diagnosis.
Case-control vs consecutive
2. Do all patients have the *reference standard*?

- Ideally all patients get the reference standard test.
- **Verification bias** = *only some* patients get the reference standard.....probably the ones in whom you really suspect have the disease.
Partial Reference Bias

Series of patients

Index test

Ref. Std. A

Compare the results of the index test with the reference standard, blinded
Differential Reference Bias

Series of patients

Index test

Ref. Std. A

Ref. Std. B

Blinded cross-classification
Series of patients

Index test

Reference standard..... includes parts of Index test

Blinded cross-classification
3. **Independent, blind or objective comparison with the reference standard?**

- Ideally, the reference standard is independent, blind and objective

- **Observer bias** = test is very subjective, or done by person who knows something about the patient or samples
Observer Bias

Series of patients

Index test

Reference standard

Unblinded cross-classification
Near patient testing for influenza in children in primary care: comparison with laboratory test

Anthony Harnden, Angela Brueggemann, Sasha Shepperd, Judy White, Andrew C Hayward, Maria Zambon, Derrick Crook, David Mant

Influenza is an important cause of acute respiratory illness in young children. Common complications include febrile convulsions, otitis media, bronchiolitis, and croup. In epidemic years attack rates among preschool children often exceed 40%. During these years children with influenza may account for up to 30% of the increase in antibiotic prescribing. Symptoms and signs of influenza in children are not specific and can mimic a range of other common respiratory viral pathogens. One quick way of reaching a precise diagnosis in primary care is to use a near

<table>
<thead>
<tr>
<th>RT-PCR test</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near patient test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>27</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>Negative</td>
<td>94</td>
<td>93</td>
<td>127</td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>96</td>
<td>157</td>
</tr>
</tbody>
</table>
Participants, methods, and results
From January to March 2001 and October to March 2002 we asked general practitioners in Oxfordshire to identify children with cough and fever who they thought had more than a simple cold. Using a nasal swab we performed a near patient test for influenza (QuickVue; Quidel, San Diego, CA). A research nurse did the test, which took 12 minutes.

We collected a nasopharyngeal aspirate from the other nostril and transported the sample to the laboratory within four hours. The laboratory staff were blind to the result of the near patient test. After adding phosphate buffered saline to the aspirate we added the emulsified sample to viral lysis buffer before freezing it at −80°C. We used RT-PCR to convert the extracted nucleic acids from RNA to complementary DNA. We performed a multiplex, nested PCR assay, using primer sets specific to influenza A and B, on all the samples. To validate our results we included quantified tissue culture specimens of influenza A and B as positive controls and water as negative control with every batch of samples tested.

A nasal swab and a nasopharyngeal aspirate were taken from 157 children. The children’s median age was 3 years (range 6 months to 12 years), and 100 were boys. We detected influenza by RT-PCR in 61 children
Teaching tips....
Diagnostic tests: What you need to know

- Validity of a diagnostic test
- Interpret the results

Set the scene, create a relaxed atmosphere;
Humour
Diagnostic Accuracy Studies

Series of patients

Index test

Reference standard

Compare the results of the index test with the reference standard, blinded

Don’t use “gold” standard

Series of patients

In pictures

Ref. Std. A

Ref. Std. B

Blinded cross-classification

CEBM

UNIVERSITY OF OXFORD
Case-control vs consecutive

Use analogies that are not medical

Get tips from other teachers!
Diagnostic Study Example

Near patient testing with reverse transcription polymerase chain reaction (RT-PCR) testing for influenza in children under 5 years of age: systematic review and meta-analysis.

Anthony Harnden, Andrew C Hayward, Maria Zambon, Judy White, Sasha Steppert, University of Oxford, Oxford and Public Health England

Influenza is a common viral illness in young children, and in children under the age of 5 years, who are otherwise healthy, it can cause fever, cough, and croup. In epidemic years, influenza in preschool children often exceeds 20% of the total contacts and 30% of the increase in antibiotic prescribing. Symptoms and signs of influenza in children are not specific and can mimic a range of other common respiratory viral pathogens. One quick way of reaching a precise diagnosis in primary care is to use a near patient test.

<table>
<thead>
<tr>
<th>RT-PCR test</th>
<th>Positive</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near patient test:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>27</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>Negative</td>
<td>34</td>
<td>93</td>
<td>127</td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>96</td>
<td>157</td>
</tr>
</tbody>
</table>

...
If you want to use something which shows potential bias, don’t use a complex test.
The Numbers
Using a brain scan, the researchers detected autism with over 90% accuracy...

You can’t diagnose autism with a brain scan...

New brain scan to diagnose autism

By Jane Hughes
Health correspondent, BBC News

A brain scan that detects autism in adults could mean much more straightforward diagnosis of the condition, scientists say.

Experts at King's College London said the scan - tested on 40 people - identified tiny but crucial signs of autism, only detectable by computer.

Current methods of diagnosis can be lengthy and expensive.

But some experts say further research will be needed before the new technique can be widely used.

The computer scan shows up a distinctive pattern associated with autism.
Appraising diagnostic tests

Are the results valid?
- Appropriate spectrum of patients?
- Does everyone get the reference standard?
- Is there an independent, blind or objective comparison with the gold standard?

What are the results?
- Sensitivity, specificity
- Likelihood ratios
- Positive and Negative Predictive Values

Will they help me look after my patients?
Pain over speed bumps in diagnosis of acute appendicitis: diagnostic accuracy study

Helen F Ashdown academic clinical fellow in general practice, Nigel D’Souza specialist registrar in general surgery, Diallah Karim foundation trainee, Richard J Stevens senior medical statistician, Andrew Huang consultant colorectal and general surgeon, Anthony Harnden university lecturer in general practice

1Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, UK; 2Department of Surgery, Stoke Mandeville Hospital, Aylesbury HP21 8AL, UK
of appendicitis. Fifty four of 64 participants were “speed bump positive.” Thirty four participants had a confirmed diagnosis of appendicitis, 33 of whom had worsened pain over speed bumps, giving a sensitivity of 97% (85% to 100%) and a specificity of 30% (15% to 49%). The positive predictive value was 61% (47% to 74%), and the negative predictive value was 90% (56% to 100%). The likelihood ratios were 1.4 (1.1 to 1.8) for a positive test result and 0.1 (0.0 to 0.7) for a negative result.
Sensitivity and Specificity
<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>True positives</td>
<td>False positives</td>
</tr>
<tr>
<td>-</td>
<td>False negatives</td>
<td>True negatives</td>
</tr>
</tbody>
</table>
### The 2 by 2 table: Sensitivity

<table>
<thead>
<tr>
<th></th>
<th>Disease</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>+</strong></td>
<td>90</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>True positives</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>False negatives</td>
</tr>
<tr>
<td><strong>-</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Proportion of people **WITH** the disease who have a **positive test result**.

Sensitivity = \( \frac{a}{a + c} \)

So, a test with 90% sensitivity.....means that the test identifies 90 out of 100 people **WITH** the disease.

Sensitivity = \( \frac{90}{100} \)
The 2 by 2 table: Specificity

<table>
<thead>
<tr>
<th>Test</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>False positives</td>
</tr>
<tr>
<td></td>
<td>25</td>
</tr>
<tr>
<td>-</td>
<td>d</td>
</tr>
<tr>
<td></td>
<td>True negatives</td>
</tr>
<tr>
<td></td>
<td>75</td>
</tr>
</tbody>
</table>

Specificity = \( \frac{d}{b + d} \)

Proportion of people **WITHOUT** the disease who have a **negative test result**.

So, a test with 75% specificity will be NEGATIVE in 75 out of 100 people **WITHOUT** the disease.

Specificity = 75/100
The Speed bump Example

Disease: Appendicitis

Test: Pain over speed bump

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>33</td>
<td>21</td>
</tr>
<tr>
<td>-</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

There were 34 people who had appendicitis... the speed bump test was positive in 33 of them.

There were 30 people who did not have appendicitis... the speed bump test was negative in 9 of them.

Sensitivity = 33/34 = 0.97 (97%)

Specificity = 9/30 = 0.30 (30%)
- Sensitivity is useful to me
  - ‘The new speed bump test was positive in 33 out of 34 people with appendicitis (sensitivity = 97%)’

- Specificity seems a bit confusing!
  - ‘The new speed bump test was negative in 9 of the 30 people who did not have appendicitis (specificity = 30%)’

- So…the false positive rate is sometimes easier
  - ‘There were 30 people who did not have appendicitis… the speed bump test was falsely positive in 21 of them’
  - So a specificity of 30% means that the new rapid test is wrong (or falsely positive) in 70% of people

Tip

True positive rate = specificity

False positive rate = 1 - specificity
Ruling In and Ruling Out

**High Sensitivity**
A good test to help in **Ruling Out** disease

High sensitivity means there are very few false negatives – so if the test comes back negative it’s highly unlikely the person has the disease.

**High Specificity**
A good test to help in **Ruling In** disease

High specificity means there are very few false positives – so if the test comes back positive it’s highly likely the person has the disease.

### Disease: Appendicitis

<table>
<thead>
<tr>
<th></th>
<th>Disease: Appendicitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test: Pain over speed bump</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>33</td>
</tr>
<tr>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

- Sensitivity = 97%
- Specificity = 30%

### Test

- Sensitivity = \( \frac{a}{a+c} \)
- Specificity = \( \frac{d}{b+d} \)
Predictive Values
Positive and Negative Predictive Value

**Disease**

Test

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>True positives</td>
<td>False positives</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>-</td>
<td>False negatives</td>
<td>True negatives</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

**PPV** = Proportion of people with a **positive test** who **have** the disease.

**PPV** = \( \frac{a}{a + b} \)

**NPV** = Proportion of people with a **negative test** who **do not** have the disease.

**NPV** = \( \frac{d}{c + d} \)

**True positives** = a

**False positives** = b

**False negatives** = c

**True negatives** = d
The Speed bump Example

Disease: Appendicitis

Test: Pain over speedbump

\[
\begin{array}{ccc}
+ & - \\
+ & 33 & 21 \\
- & 1 & 9 \\
\end{array}
\]

- PPV = 33/54 = 61%
- NPV = 9/10 = 90%
Your father went to his doctor and was told that his test for a disease was positive. He is really worried, and comes to ask you for help!

After doing some reading, you find that for men of his age:
The prevalence of the disease is 30%
The test has sensitivity of 50% and specificity of 90%

“Tell me what’s the chance I have this disease?”
Disease has a prevalence of 30%. The test has sensitivity of 50% and specificity of 90%.

<table>
<thead>
<tr>
<th>Predictive Value</th>
<th>100%</th>
<th>Likely</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>Maybe</td>
<td></td>
</tr>
<tr>
<td>0%</td>
<td>Unlikely</td>
<td></td>
</tr>
</tbody>
</table>
Disease has a prevalence of 30%.
The test has sensitivity of 50% and specificity of 90%.
Given a positive test, what is the probability your dad has the disease?
Prevalence of 30%, Sensitivity of 50%, Specificity of 90%

22 people test positive........ of whom 15 have the disease

So, chance of disease is $\frac{15}{22} = 68\%$
Prevalence of 4%, Sensitivity of 50%, Specificity of 90%

- Disease +ve
  - 4 people test positive
  - 2 of whom have the disease
  - Chance of disease is $\frac{2}{11.6} = 17\%$

- Disease -ve
  - 96 people
  - Sensitivity = 50%
  - False positive rate = 10%
  - 11.6 people test positive
    - of whom 2 have the disease

So, chance of disease is $\frac{2}{11.6} = 17\%$
Positive and Negative Predictive Value

NOTE

• PPV and NPV are not intrinsic to the test – they also depend on the prevalence!

• NPV and PPV should only be used if the ratio of the number of patients in the disease group and the number of patients in the healthy control group is equivalent to the prevalence of the diseases in the studied population

• Use Likelihood Ratio - does not depend on prevalence
Teaching tips….
You can’t diagnose autism with a brain scan...

Using a brain scan, the researchers detected autism with over 90% accuracy...

Use examples from the news, blogs, things that people come across – relevant to everyone, not just clinicians; Suspense...

A brain scan that detects autism in adults could mean much more straightforward diagnosis of the condition, scientists say.

Experts at King’s College London said the scan, tested on 40 people — identified tiny but crucial signs of autism, only detectable by computer.

Current methods of diagnosis can be lengthy and expensive.

But some experts say further research will be needed before the new technique can be widely used.

The computed scan shows up a distinctive pattern associated with autism
Find a simple paper with different measures and the actual numbers.
The 2 by 2 table: Sensitivity

<table>
<thead>
<tr>
<th>Disease</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>90</td>
</tr>
<tr>
<td>-</td>
<td>10</td>
</tr>
</tbody>
</table>

Sensitivity = \( \frac{a}{a + c} \)

Explain the concepts in words. Don’t focus on formulas – some like them (so provide them), but for many this feels too much like “MATHS!”

Sensitivity = \( \frac{90}{100} \)
The Speed bump Example

Disease: Appendicitis

Test: Pain over speed bump

- Use numbers from a paper; simple language; It’s more important to understand what it all means than to know how to calculate

There were 34 people who had appendicitis... the speed bump test was positive in 33 of them

There were 30 people who did not have appendicitis... the speed bump test was negative in 9 of them

Sensitivity = 33/34 = 0.97 (97%)

Specificity = 9/30 = 0.30 (30%)
Tip

- **Sensitivity** is useful to me
  - ‘The new speed bump test was positive in 33 out of 34 people with appendicitis (sensitivity = 97%)’

- Specificity seems a bit confusing!
  - ‘The new speed bump test was negative in 9 of the 30 people who did not have appendicitis (specificity = 30%)’

- So…the false positive rate is sometimes easier
  - ‘There were 30 people who did not have appendicitis… the speed bump test was falsely positive in 21 of them’
  - So a specificity of 30% means that the new rapid test is wrong (or falsely positive) in 70% of people

**True positive rate = specificity**

**False positive rate is easier to understand than specificity – provide options!**
Ruling In and Ruling Out

High Sensitivity

A good test to help in Ruling In

High Specificity

A good test to help in Ruling Out

High sensitivity means there are very few false negatives – so if the test comes back negative it's highly unlikely the person has the disease.

High specificity means there are very few false positives – so if the test comes back positive it's highly likely the person has the disease.

Disease

Test

<table>
<thead>
<tr>
<th></th>
<th>Appendicitis</th>
<th>Pain over speed bump</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

Sensitivity = 97%

Specificity = 30%

Acronyms help some...but confuse others

For beginners this may be a step too far...

Acronyms help some...but confuse others

Touch on it...then park it and move on...

SnNOUT

SpPIN

Sensitivity = \( \frac{a}{a+c} \)

Specificity = \( \frac{d}{b+d} \)
Your father went to his doctor and was told that his test for a disease was positive. He is really worried, and comes to ask you for help.

After doing some reading, you find that for men of his age:

- The prevalence of the disease is 30%.
- The test has sensitivity of 50% and specificity of 90%.

"Tell me what’s the chance I have this disease?"
Disease has a prevalence of 30%.
The test has sensitivity of 50% and specificity of 90%.

Predictive Value

- 100% Likely
- 50% Maybe
- 0% Unlikely

Have a go... interactive... safe environment
Natural Frequencies

Disease has a prevalence of 30%.
The test has sensitivity of 50% and specificity of 90%.
Given a positive test, what is the probability your dad has the disease?

Set a time and stick to it!
Prevalence of 30%, Sensitivity of 50%, Specificity of 90%

Simple numbers = year 2 maths;
reinforces sensitivity and specificity;
No formulas!

22 people test positive...........
of whom 15 have the disease

So, chance of disease is
15/22 = 68%

Disease +ve

Disease -ve

100

70

7

False positive rate = 10%
Prevalence of 4%, Sensitivity of 50%, Specificity of 90%

11.6 people test positive........ of whom 2 have the disease

So, chance of disease is 2/11.6 = 17%

Change the prevalence, keep other numbers the same... learning by doing; Good transition to likelihood ratios
Likelihood Ratios
Likelihood ratios

\[
LR = \frac{\text{Probability of clinical finding in patients with disease}}{\text{Probability of same finding in patients without disease}}
\]

Example:
If 80% of people with a cold have a runny nose and 10% of people without a cold have a runny nose, then the LR for runny nose is:
80%/10% = 8
## Likelihood ratios

### Positive likelihood ratio (LR+)

How much more likely is a **positive test** to be found in a person **with the disease** than in a person **without it**?

\[
LR^+ = \frac{\text{sens}}{1 - \text{spec}}
\]

### Negative likelihood ratio (LR-)

How much more likely is a **negative test** to be found in a person **without the disease** than in a person **with it**?

\[
LR^- = \frac{1 - \text{sens}}{\text{spec}}
\]
What do likelihood ratios mean?

- LR<0.1 = strong negative test result
- LR=1 = No diagnostic value
- LR>10 = strong positive test result
Diagnosis of Appendicitis

McBurney’s point

Rovsing’s sign
If palpation of the left lower quadrant of a person's abdomen results in more pain in the right lower quadrant

Psoas sign
Abdominal pain resulting from passively extending the thigh of a patient or asking the patient to actively flex his thigh at the hip

Ashdown’s sign
Pain when driving over speed bumps
For Example

Speed bump test (Ashdown’s sign):
LR+ = 1.4
LR- = 0.1
Bayesian reasoning

Pre test 5%

?Appendicitis:

McBurney tenderness LR+ = 3.4

Speed bump test LR- = 0.1

Post-test odds = Pre-test odds x Likelihood ratio

Fagan nomogram

Post-test odds for disease after one test become pre-test odds for next test etc.

Post test ~20%

Post test ~0.5%
Teaching tips....
Likelihood ratios

\[ LR = \frac{\text{Probability of clinical finding in patients with disease}}{\text{Probability of same finding in patients without disease}} \]

Example:
If 80% of people with a cold have a runny nose and 10% of people without a cold have a runny nose, then the LR for runny nose is:
80%/10% = 8

Definition is wordy...so give a simple example
Positive likelihood ratio (LR+): How much more likely is a positive test to be found in a person with the disease than in a person without it?

\[ LR^+ = \frac{\text{sens}}{1 - \text{spec}} \]

Calculation in terms of sensitivity/ specificity is simpler and more useful than formula from the 2x2 table.

Negative likelihood ratio (LR-): How much more likely is a negative test to be found in a person without the disease than in a person with it?

\[ LR^- = \frac{1 - \text{sens}}{\text{spec}} \]
What do likelihood ratios mean?

Knowing what LRs mean is more important than how to calculate.

- **LR<0.1** = strong negative test result
- **LR=1** = No diagnostic value
- **LR>10** = strong positive test result
Diagnosis of Appendicitis

**McBurney’s point**

**Rovsing’s sign**

If palpation of the left lower quadrant of a person’s abdomen results in more pain in the right lower quadrant, it is a sign of appendicitis.

**Ashdown’s sign**

Pain when driving over speed bumps

Simple example... related to the paper
For Example

Speed bump test (Ashdown’s sign):

- LR+ = 1.4
- LR- = 0.1

Putting numbers on the scale makes it clearer

LR+ = 3.4
LR- = 0.4
Post test ~20%

?Appendicitis:
McBurney tenderness

Speed bump test LR- = 0.1

Post-test odds = Pre-test odds x Likelihood ratio

Key concept: Nomogram links pre- and post-test odds; Keep it to a minimum

Post-test odds for disease after one test become pre-test odds for next test etc.

Bayesian reasoning

Fagan nomogram
What about the news story...?
A brain scan that detects autism in adults could mean much more straightforward diagnosis of the condition, scientists say.

Experts at King's College London said the scan - tested on 40 people - identified tiny but crucial signs of autism, only detectable by computer.

Current methods of diagnosis can be lengthy and expensive.

But some experts say further research will be needed before the new technique can be widely used.

The researchers detected autism with over 90% accuracy, the Journal of Neuroscience reports.
Describing the Brain in Autism in Five Dimensions—Magnetic Resonance Imaging-Assisted Diagnosis of Autism Spectrum Disorder Using a Multiparameter Classification Approach

Christine Ecker,1 Andre Marquand,2 Janaina Mourão-Miranda,3,4 Patrick Johnston,1 Eileen M. Daly,1 Michael J. Brammer,2 Stefanos Maltezos,1 Clodagh M. Murphy,1 Dene Robertson,1 Steven C. Williams,3 and Declan G. M. Murphy1

1Section of Brain Maturation, Department of Psychological Medicine, Institute of Psychiatry, 2Brain Image Analysis Unit, Department of Biostatistics, Institute of Psychiatry, and 3Centre for Neuroimaging Sciences, Institute of Psychiatry, King's College, London SE5 8AF, United Kingdom, and 4Centre for Computational Statistics and Machine Learning, Department of Computer Science, University College London, London WC1E 6BT, United Kingdom

Table 3. Results of SVM classification between ASD and control group using different brain morphometric features in the left and right hemispheres

<table>
<thead>
<tr>
<th>Morphometric feature</th>
<th>Correctly classified (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Left hemisphere</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All parameters</td>
<td>85</td>
<td>90</td>
<td>80</td>
<td>0*</td>
</tr>
<tr>
<td>Cortical thickness</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Radial curvature</td>
<td>72.5</td>
<td>65</td>
<td>80</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average convexity</td>
<td>70</td>
<td>75</td>
<td>65</td>
<td>&lt;0.004</td>
</tr>
<tr>
<td>Metric distortion</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>0*</td>
</tr>
<tr>
<td>Plai area</td>
<td>77.5</td>
<td>70</td>
<td>85</td>
<td>0*</td>
</tr>
<tr>
<td><strong>Right hemisphere</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All parameters</td>
<td>65</td>
<td>60</td>
<td>70</td>
<td>&lt;0.03</td>
</tr>
<tr>
<td>Cortical thickness</td>
<td>60</td>
<td>65</td>
<td>55</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Radial curvature</td>
<td>52.5</td>
<td>50</td>
<td>55</td>
<td>&lt;0.30</td>
</tr>
<tr>
<td>Average convexity</td>
<td>50</td>
<td>40</td>
<td>60</td>
<td>&lt;0.40</td>
</tr>
<tr>
<td>Metric distortion</td>
<td>57.5</td>
<td>45</td>
<td>70</td>
<td>&lt;0.06</td>
</tr>
<tr>
<td>Plai area</td>
<td>45</td>
<td>45</td>
<td>45</td>
<td>&lt;0.60</td>
</tr>
</tbody>
</table>

Correctly identified ASD cases were considered true positive. *p values of zero indicate that not a single one of the 1000 permutations provided a better classification.

The indication from recent studies is that the figures cannot be precisely fixed, but it appears that a prevalence rate of around 1% is a best estimate a best estimate of the prevalence in children. No prevalence studies have ever been carried out on adults.
Autism has a prevalence of 1%.
The test has sensitivity of 90% and specificity of 80%.
Given a positive test, what is the probability the child has autism?
Prevalence of 1%, Sensitivity of 90%, Specificity of 80%

- Disease +ve
  - 100 people
  - 1 disease positive
  - 99 disease negative

- Disease -ve
  - Testing +ve
  - Sensitivity = 90%
  - False positive rate = 20%
  - 0.9 people test positive
  - 19.8 people test negative

- Testing +ve
  - 20.7 people test positive
  - Of whom 0.9 have the disease

So, chance of disease is $0.9/20.7 \approx 4.5\%$
autism and brain scan test: the real

What has happened is the sensitivity has been taken for the positive predictive value, which is what you want to know: if I have a positive test do I have the disease?

Sensitivity: The proportion of people with disease who have a positive test. Positive predictive value (+PV): The proportion of people with a positive test who have disease.

So, for a prevalence of 1%, the actual positive predictive value is 4.5%. That is about 5 in every 100 with a positive test would have autism. Even at a prevalence of 2%, only 8.5% would be correctly identified.

Suddenly, not that great a test. This has to be one of the worst examples of misinterpreting diagnostic test results in the media I've ever seen.
Why autism can't be diagnosed with brain scans

Using brain scans to detect autism would be a huge expensive waste of money, says Carl Heneghan

The BBC, the Guardian and Reuters this week widely reported British researchers published in the Journal of Neuroscience have developed a brain scan which can detect autism in adults with 90% accuracy.

Dr Christine Ecker, the lead author, showed her imaging technique was able to detect which people in her group had autism. "If we get a new case, we will also hopefully be 90% accurate," she said.

Pretty simple then, you turn up, have the test, and you have a 90% chance of finding out whether you have autism.

Well, you couldn't be any further from the truth.
Stomach cancer 'spotted by breath test'

By Michelle Roberts
Health editor, BBC News online

A quick and simple breath test can diagnose stomach cancer, study findings reveal.

Scientists from Israel and China found the test was 90% accurate at detecting and distinguishing cancers from other stomach complaints in 130 patients.

The British Journal of Cancer says the test could revolutionise and speed up the way this cancer is diagnosed.

About 7,000 UK people develop stomach cancer each year and most have an advanced stage of the disease.

Two-fifths of patients survive for at least a year, but only a fifth are still alive after five years, despite treatment.

Currently doctors diagnose stomach cancer by taking a biopsy of the stomach lining using a probe and a flexible camera passed via mouth and down the gullet.

The new test looks for chemical profiles in exhaled breath that are unique to patients with stomach cancer.

Volatile organic compounds

Cancer appears to give off a signature smell of volatile organic compounds that can be detected using the right technical medical kit - and perhaps even dogs.

The science behind the test itself is not new - many researchers have been working on the possibility of breath tests for a number of cancers.
ARE YOU COMING TO BED?

I CAN'T. THIS IS IMPORTANT.

WHAT?

SOMEONE IS WRONG ON THE INTERNET.
Appraising diagnostic tests

Are the results valid?

- Appropriate spectrum of patients?
- Does everyone get the gold standard?
- Is there an independent, blind or objective comparison with the gold standard?

What are the results?

- Sensitivity, specificity
- Likelihood ratios
- Positive and Negative Predictive Values

Will they help me look after my patients?

- Can I do the test in my setting?
- Do results apply to the mix of patients I see?
- Will the result change my management?
- Costs to patient/health service?
<table>
<thead>
<tr>
<th>Will the test apply in my setting?</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Reproducibility of the test and interpretation in my setting</td>
</tr>
<tr>
<td>• Do results apply to the mix of patients I see?</td>
</tr>
<tr>
<td>• Will the results change my management?</td>
</tr>
<tr>
<td>• Impact on outcomes that are important to patients?</td>
</tr>
<tr>
<td>• Where does the test fit into the diagnostic strategy?</td>
</tr>
<tr>
<td>• Costs to patient/health service?</td>
</tr>
</tbody>
</table>
What is the ONE thing I need to remember from today?

Are the results valid?

What are the results?

Will they help me look after my patients?

Don’t believe everything you are told, Ask for the Evidence!
Teaching tips….
The researchers detected autism with over 90% accuracy, the Journal of Neuroscience reports.
Will the test apply in my setting?

- Reproducibility of the test and its interpretation in my setting?
- Do results apply to the mix of patients I see?
- Will the results change my management?
- Impact on outcomes that are important to patients?
- Where does the test fit into the diagnostic strategy?
- Costs to patient/health service?

There is more to diagnostics than accuracy!
Are the results valid?

Take home message!

Will they help me look after my patients?

Don’t believe everything you are told, Ask for the Evidence!
<table>
<thead>
<tr>
<th>Useful books on diagnostics</th>
</tr>
</thead>
</table>
| **Evidence base of Clinical Diagnosis.**  
Knottnerus & Buntinx.  
Wiley-Blackwell |
| **Evidence-based Diagnosis.**  
Newman & Kohn.  
Cambridge Univ. Press |
| **Diagnostic Tests Toolkit.**  
Thompson & Van den Bruel.  
Wiley-Blackwell. |
| **Evidence based Physical Diagnosis.**  
Steven McGee.  
Saunders |
| **The Diagnostic Process.**  
John Balla.  
Cambridge Univ. Press |
## Useful journal articles on diagnostics

- Rutjes et al. Evidence of bias and variation in diagnostic accuracy studies. CMAJ 2006