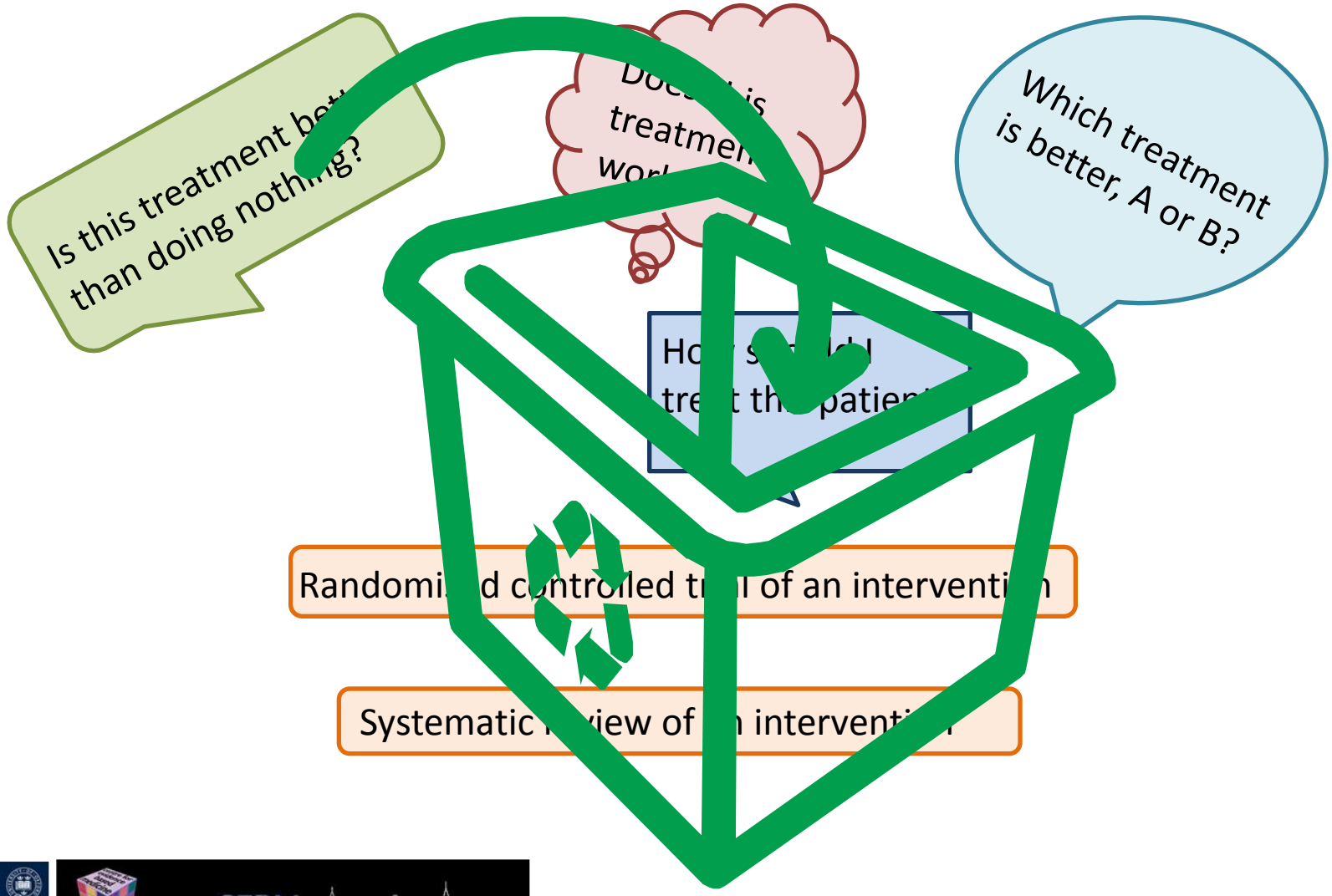
Diagnostic Studies

Dr. Annette Plüddemann

Department of Primary Care Health Sciences, University of Oxford
Centre for Evidence Based Medicine

What kinds of EBM questions have you asked?



Diagnostic studies: What you need to know

- Validity of a diagnostic study
- Interpret the results

"Mr. Osborne, may I be excused? My brain is full."

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

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

BBC Mobile | News | Sport | Weather | iPlayer | TV

NEWS HEALTH

Home | World | UK | England | N. Ireland | Scotland | Wales | Business | Politics | Health | Education | Sci/Envir

10 August 2010 Last updated at 22:01

New brain scan to diagnose autism

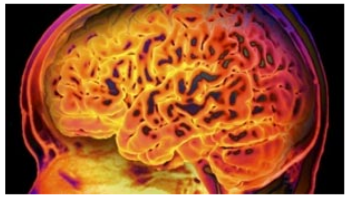
By Jane Hughes
Health correspondent, BBC News

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

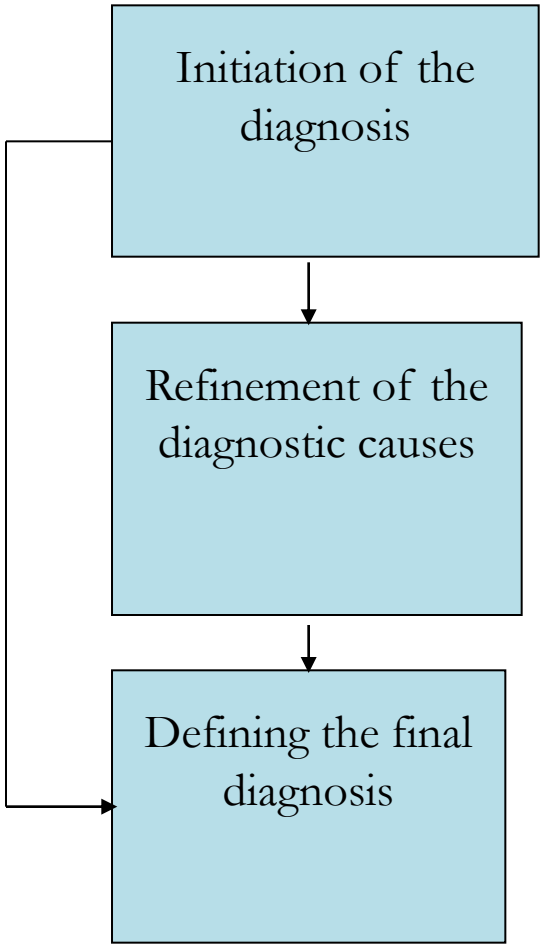
How do clinicians make diagnoses?

- Patient history…examination…differential diagnosis…final diagnosis

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

(Diagnostic strategies used in primary care. Heneghan, et al,. *BMJ 2009. 20;338:b9462*009)

Diagnostic stages & strategies



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

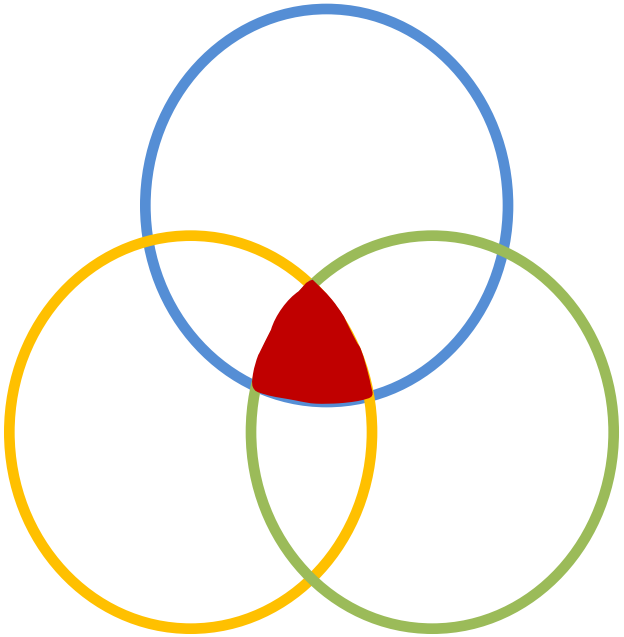
Initiation: Self-labelling

- 20% of consultations
- Accuracy of self-diagnosis in recurrent UTI
  - 88 women with 172 self-diagnosed UTIs
    - Uropathogen in 144 (84%)
    - Sterile pyuria in 19 cases (11%)
    - No pyuria or bacteriuira in 9 cases (5%)

(Gupta *et al.* Ann Int Med 2001)

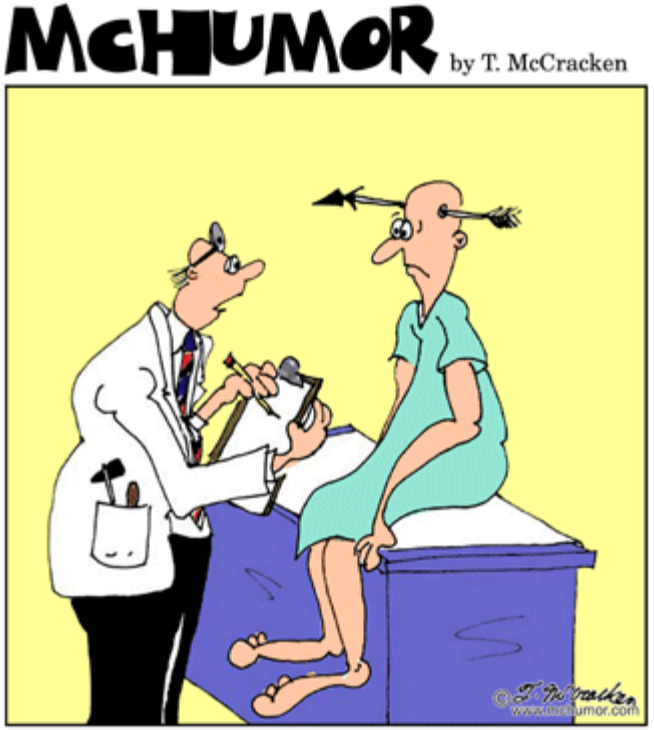
Diagnostic reasoning

- Pattern recognition
- Rule out
- Prediction rules
- Test hypothesis
- Red flags
- Response to a therapy
- Time
- Rules of thumb ‘Heuristics’



What are tests used for?

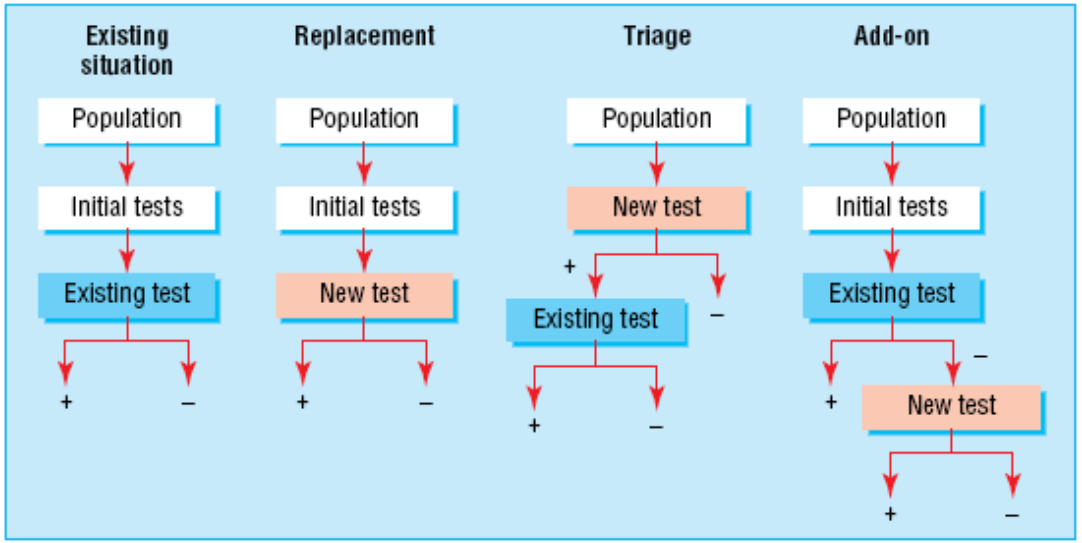
- Increase certainty about presence/absence of disease
- Disease severity
- Monitor clinical course
- Assess prognosis – risk/stage within diagnosis
- Plan treatment e.g., location
- Stall for time!



"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



Roles of tests and positions in existing diagnostic pathways

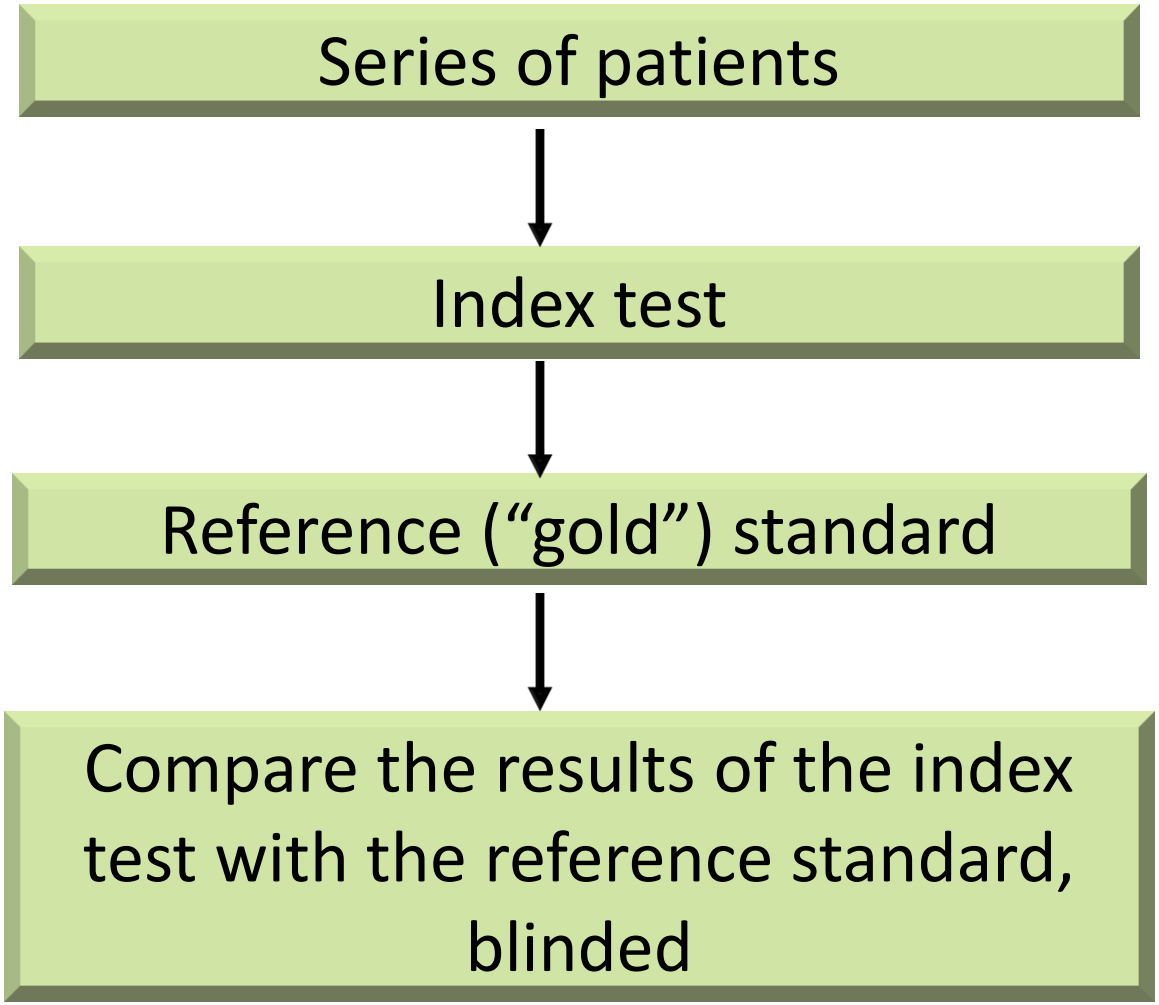
Bossuyt *et al.* BMJ 2006;332:1089–92

Interpreting Diagnostic Studies

Is this study valid?

What do all the numbers mean??

Diagnostic Studies



Diagnostic Study Example

Primary care

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

Department of Primary Health Care, Institute of Health Sciences, University of Oxford, Oxford OX3 7LF

Anthony Harnden *university lecturer*

Sasha Shepperd *university research lecturer*

Judy White *research nurse*

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.[1] 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

Comparison of near patient testing with reverse transcription polymerase chain reaction (RT-PCR) testing for influenza in children

	RT-PCR test		
	Positive	Negative	Total
Near patient test:			
Positive	27	3	30
Negative	34	93	127
Total	61	96	157

BMJ VOLUME 326 MARCH 2003

Appraising diagnostic studies: 3 easy steps

Are the results valid?

What are the results?

Will they help me look after my patients?

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

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

1. Spectrum

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

2. Do all patients have the ***gold standard***?

- Ideally all patients get the gold /reference standard test

1. Spectrum

2. Index test

3. Gold standard

Participants, methods, and results

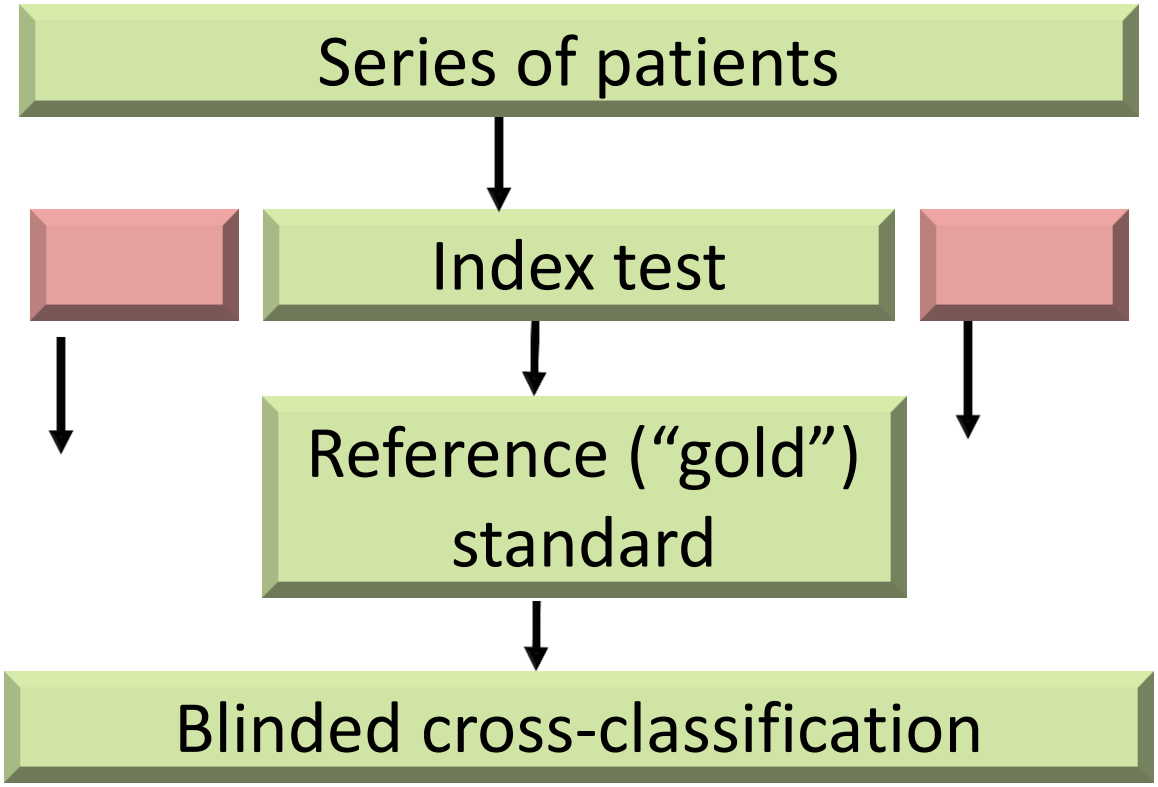
From January to March 2001 and October to March 2002 we asked general practitioners in Oxfordshire to identify children with cough and fever who they thought had more than a simple cold. Using a nasal swab we performed a near patient test for influenza (QuickVue; Quidel, San Diego, CA). A research nurse 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.

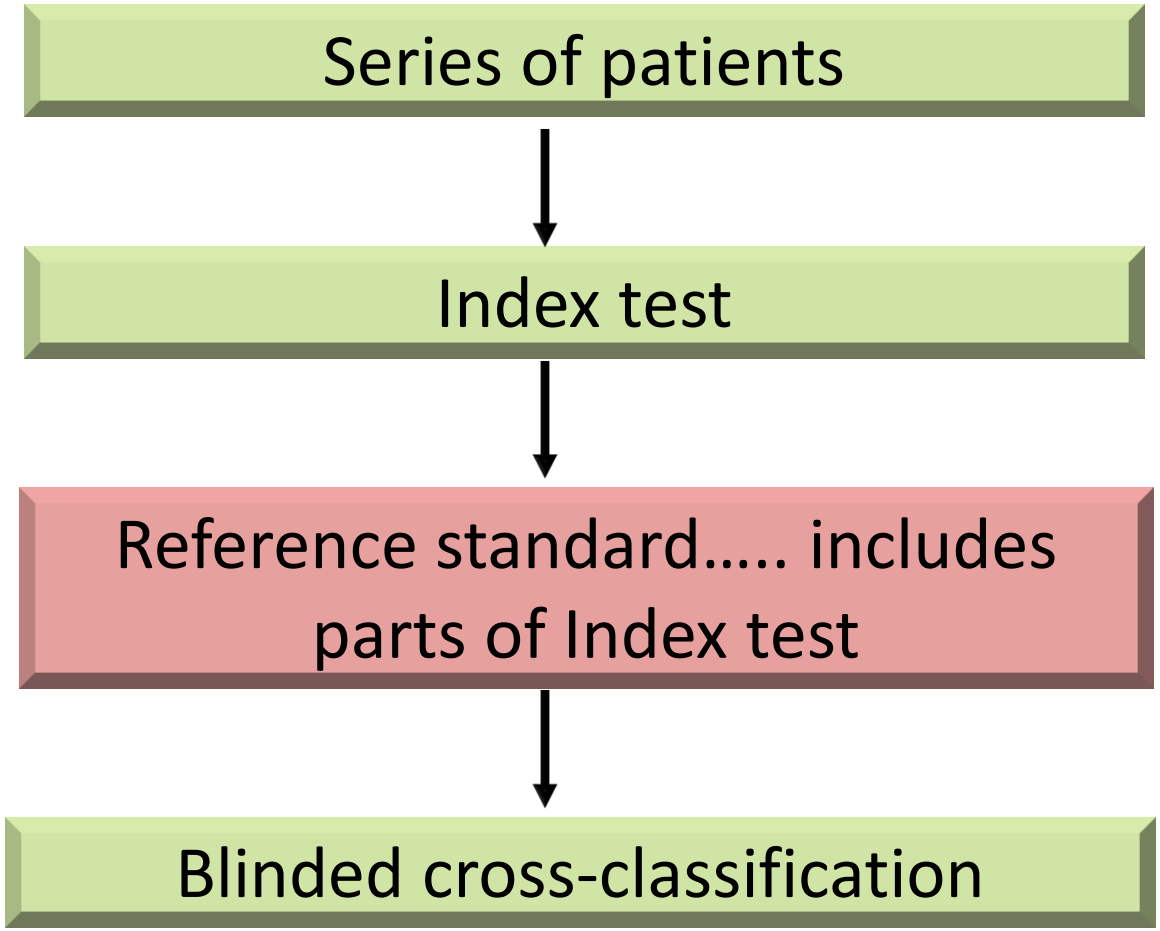
A nasal swab and a nasopharyngeal aspirate were taken from 157 children. The children's median age was 3 years (range 6 months to 12 years), and 100 were boys. We detected influenza by RT-PCR in 61 children

Verification (work-up) Bias

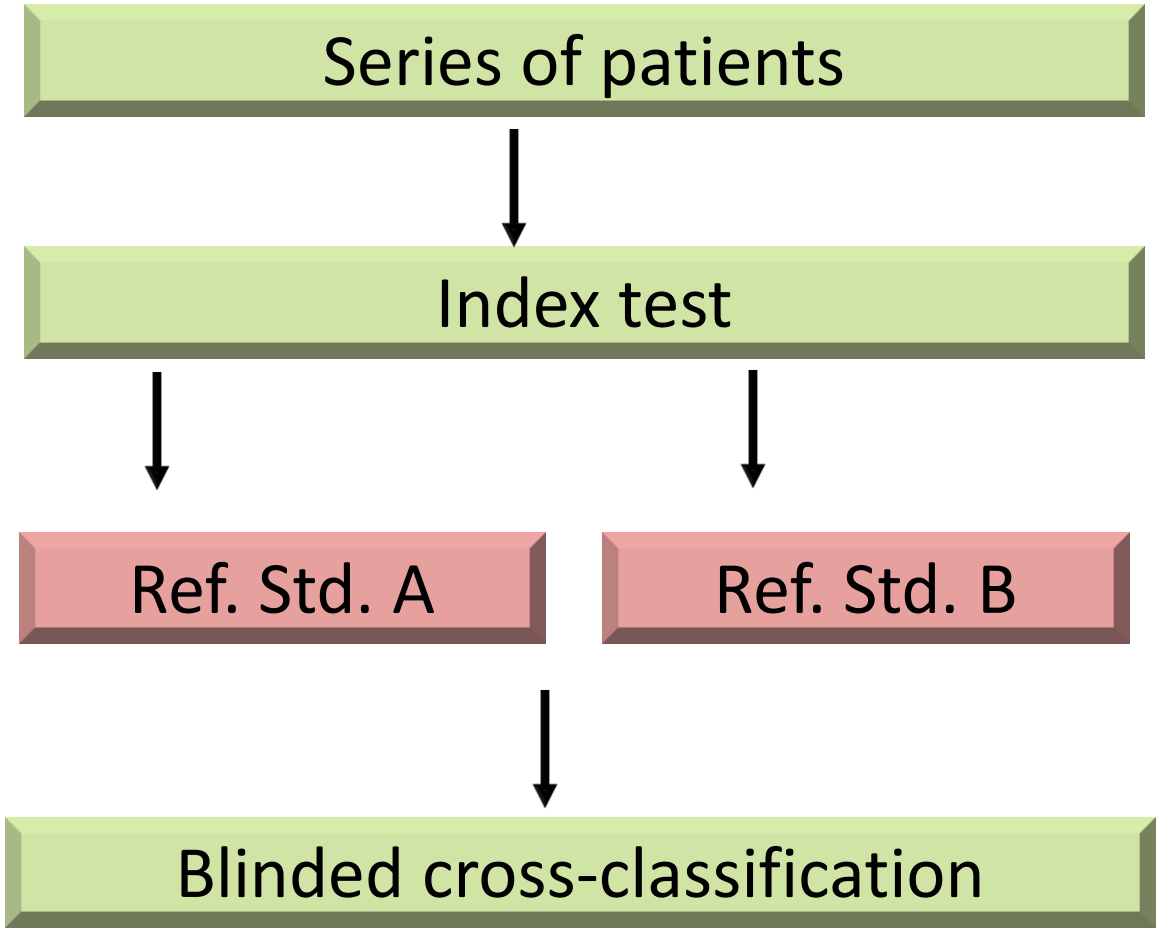
Only **some** patients get the gold standard…..probably the ones in whom you really suspect have the disease



Incorporation Bias



Differential Reference Bias



3. ***Independent, blind or objective comparison*** with the gold standard?

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

1. Spectrum

2. Index test

3. Gold standard

4. Blinding

Participants, methods, and results

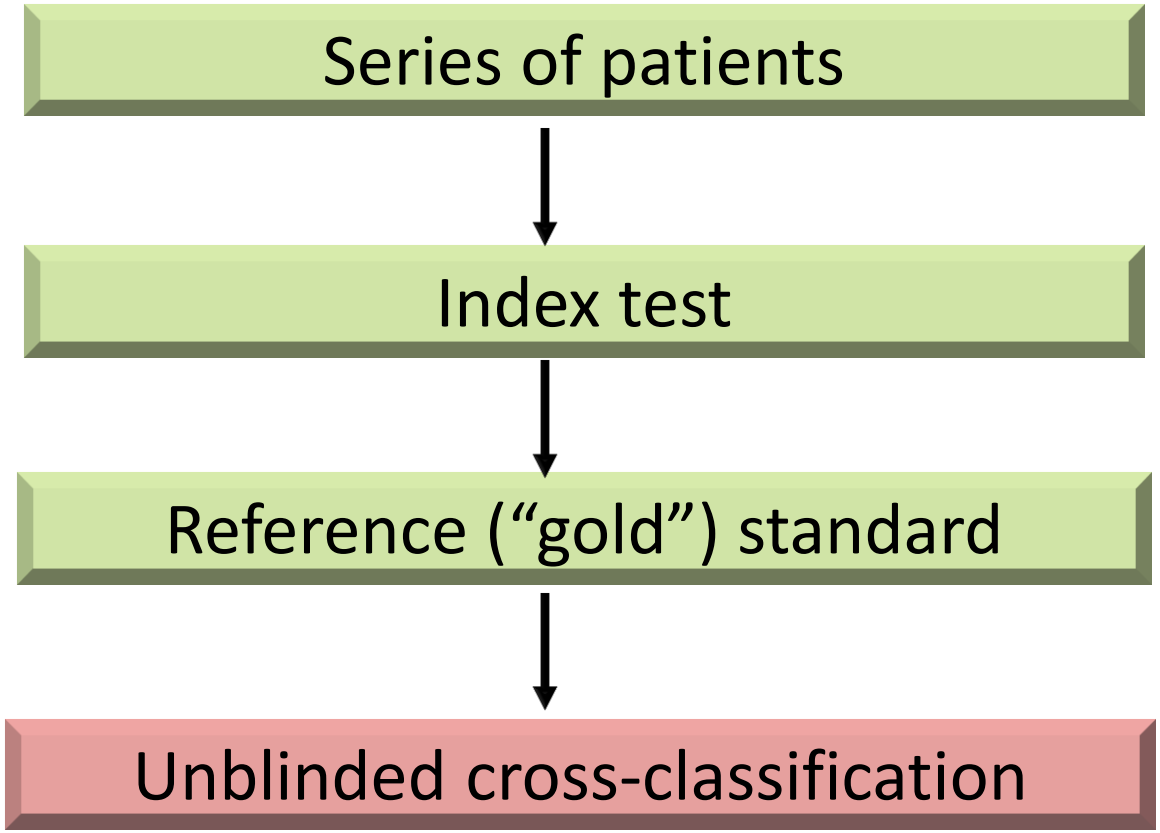
From January to March 2001 and October to March 2002 we asked general practitioners in Oxfordshire to identify children with cough and fever who they thought had more than a simple cold. Using a nasal swab we performed a near patient test for influenza (QuickVue; Quidel, San Diego, CA). A research nurse 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

Observer Bias

Test is very subjective, or done by person who knows something about the patient or samples



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?

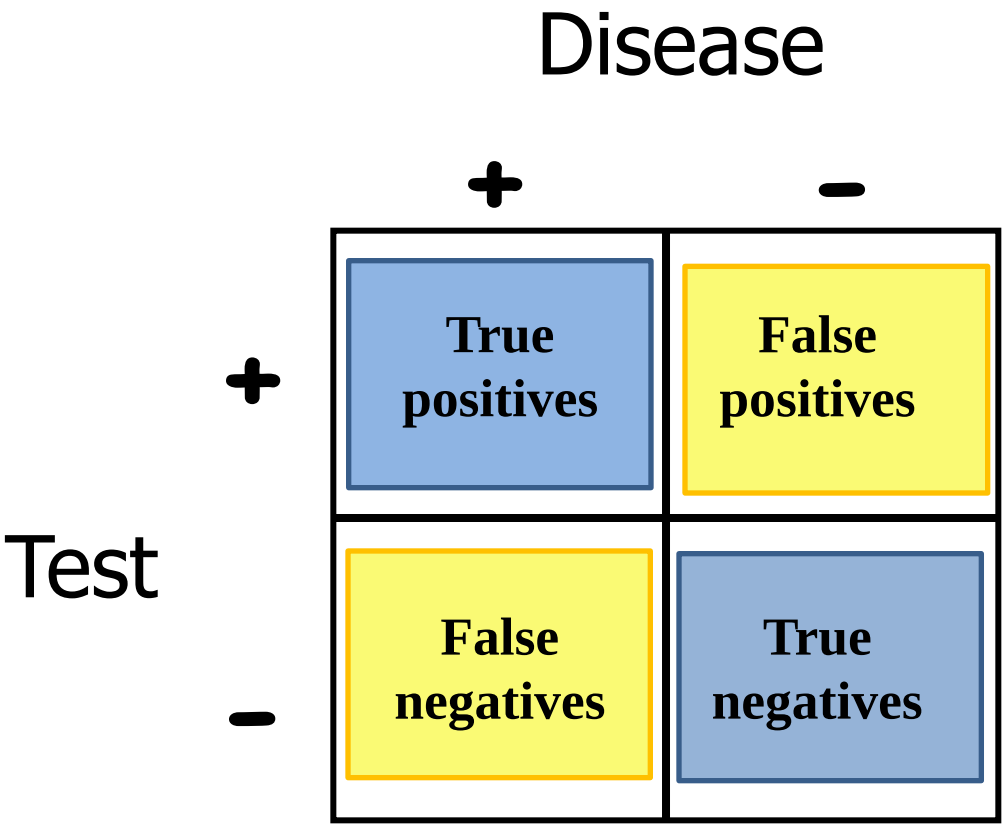
What are the results?

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

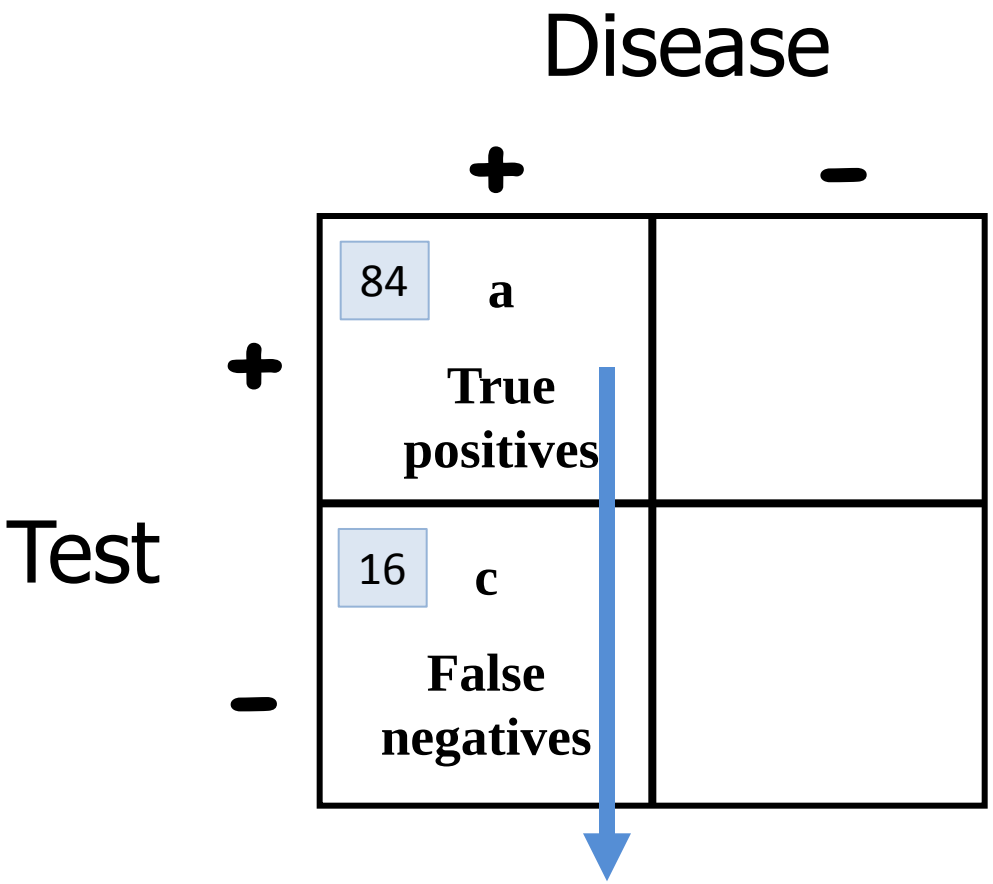
Will they help me look after my patients?

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 (39%). The near patient test was positive in 27 of these 61 children, giving a sensitivity of 44% (95% confidence interval 32% to 58%) and a specificity of 97% (91% to 99%) (table). The likelihood ratio for a positive test result was 14.2 (4.5 to 44.7) and for a negative result 0.58 (0.46 to 0.72).

The 2 by 2 table



The 2 by 2 table: Sensitivity



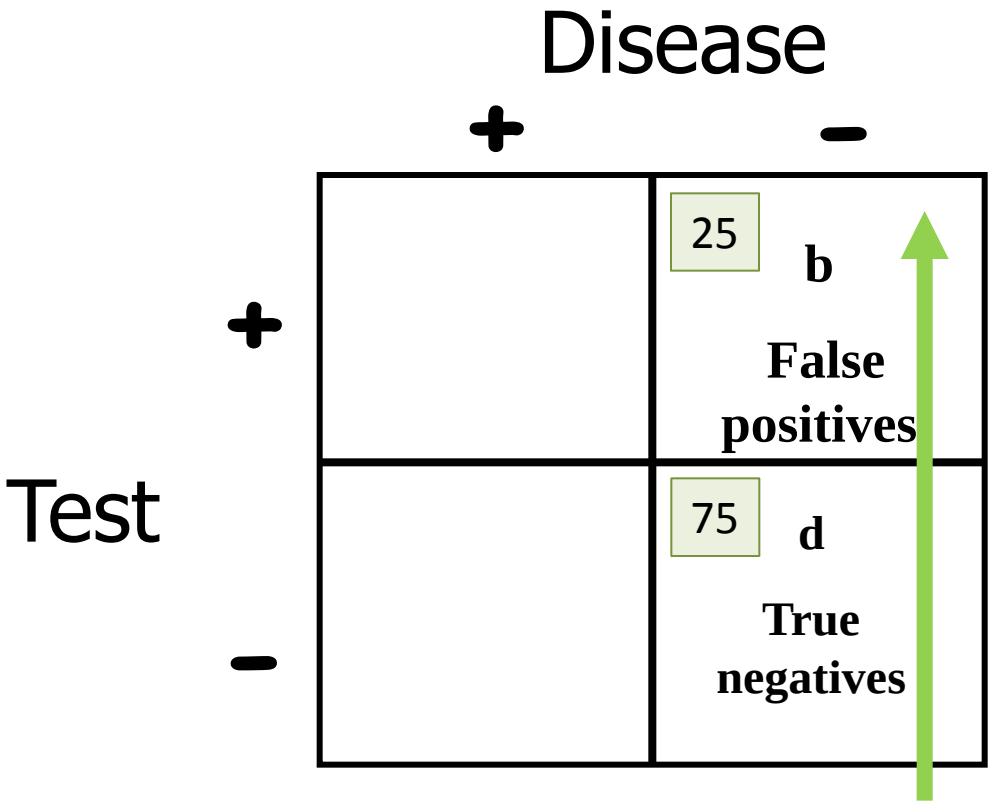
Proportion of people WITH the disease who have a positive test result.

So, a test with 84% sensitivity….means that the test identifies 84 out of 100 people WITH the disease

Sensitivity = a / a + c

Sensitivity = 84/100

The 2 by 2 table: Specificity



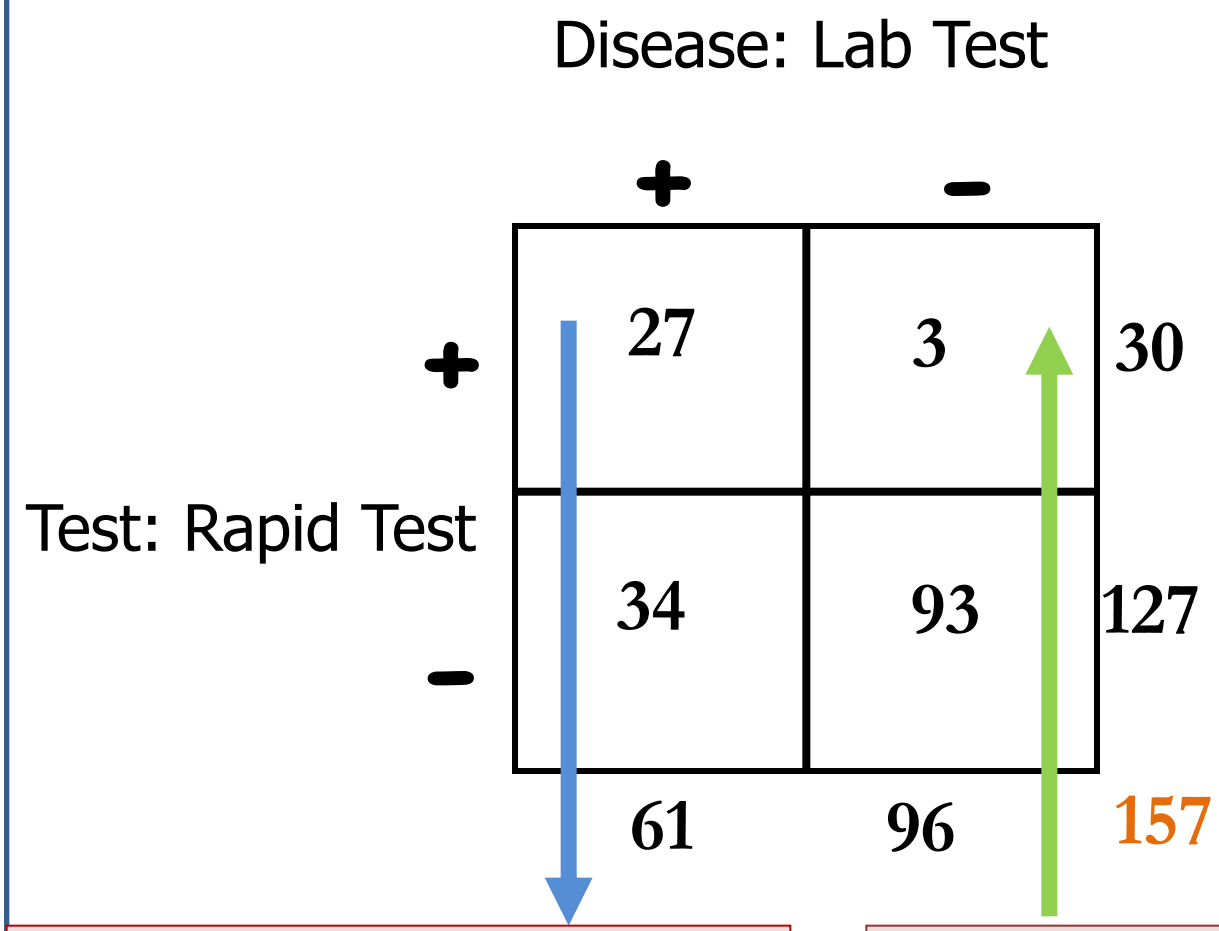
Proportion of people WITHOUT the disease who have a negative test result.

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

Specificity = d / b + d

Specificity = 75/100

The Influenza Example



There were 61 children who had influenza…the rapid test was positive in 27 of them

There were 96 children who did not have influenza… the rapid test was negative in 93 of them

Sensitivity = 27/61 = 0.44 (44%)

Specificity = 93/96 = 0.97 (97%)

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 (39%). The near patient test was positive in 27 of these 61 children, giving a sensitivity of 44% (95% confidence interval 32% to 58%) and a specificity of 97% (91% to 99%) (table). The likelihood ratio for a positive test result was 14.2 (4.5 to 44.7) and for a negative result 0.58 (0.46 to 0.72).

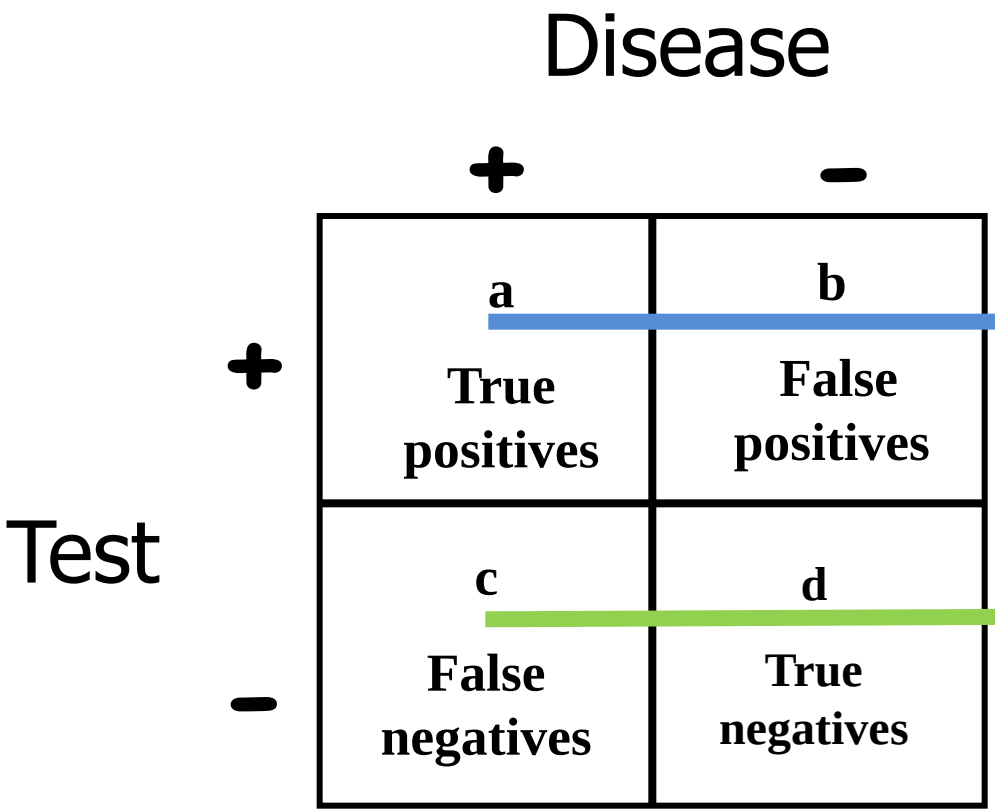
Tip

- **Sensitivity** is useful to me
  - 'The new rapid influenza test was positive in 27 out of 61 children with influenza (sensitivity = 44%)'
- Specificity seems a bit confusing!
  - 'The new rapid influenza test was negative in 93 of the 96 children who did not have influenza (specificity = 97%)'
- So…the **false positive rate** is sometimes easier

  False positive rate = 1 - specificity

  - 'There were 96 children who did not have influenza… the rapid test was falsely positive in 3 of them'
  - So a specificity of 97% means that the new rapid test is wrong (or falsely positive) in 3% of children

Positive and Negative Predictive Value



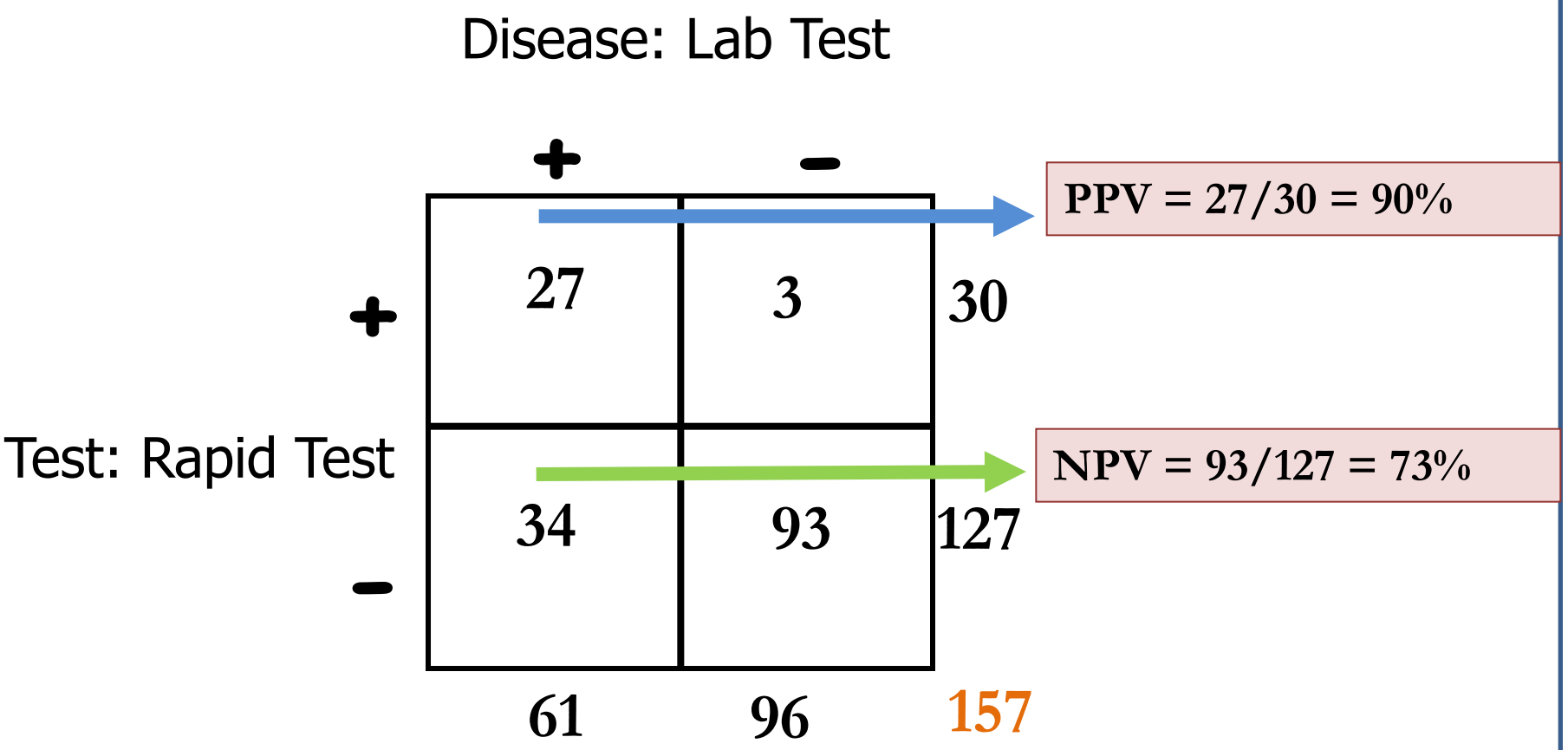
PPV = Proportion of people with a positive test who have the disease.

PPV = a / a + b

NPV = d / c + d

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

The Influenza Example



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 disease in the studied population

- Use Likelihood Ratio - does not depend on prevalence

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?

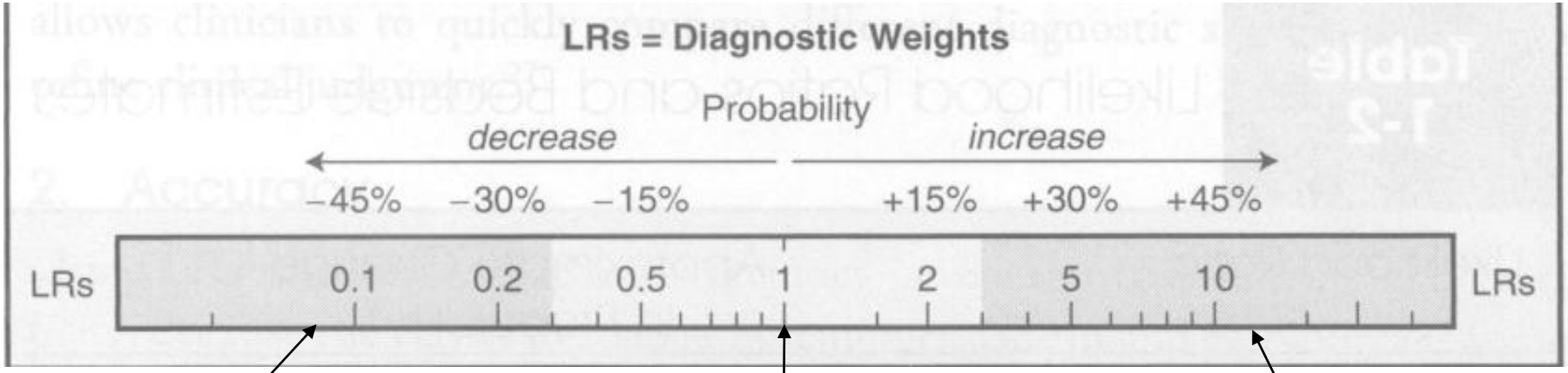
LR+ = sens/(1-spec)

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- = (1-sens)/(spec)

What do likelihood ratios mean?



LR<0.1 = strong negative test result

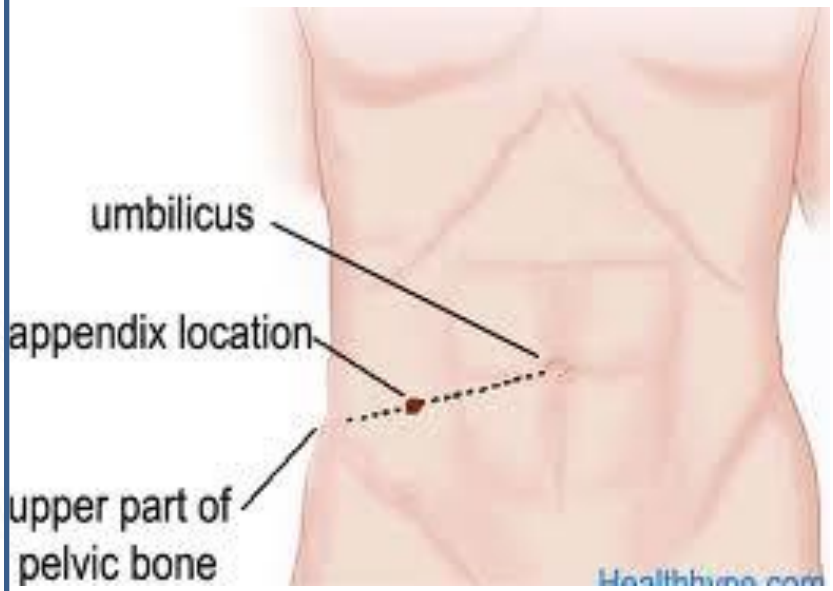
LR=1

No diagnostic value

LR>10 = strong positive test result

Diagnosis of Appendicitis

McBurney's point



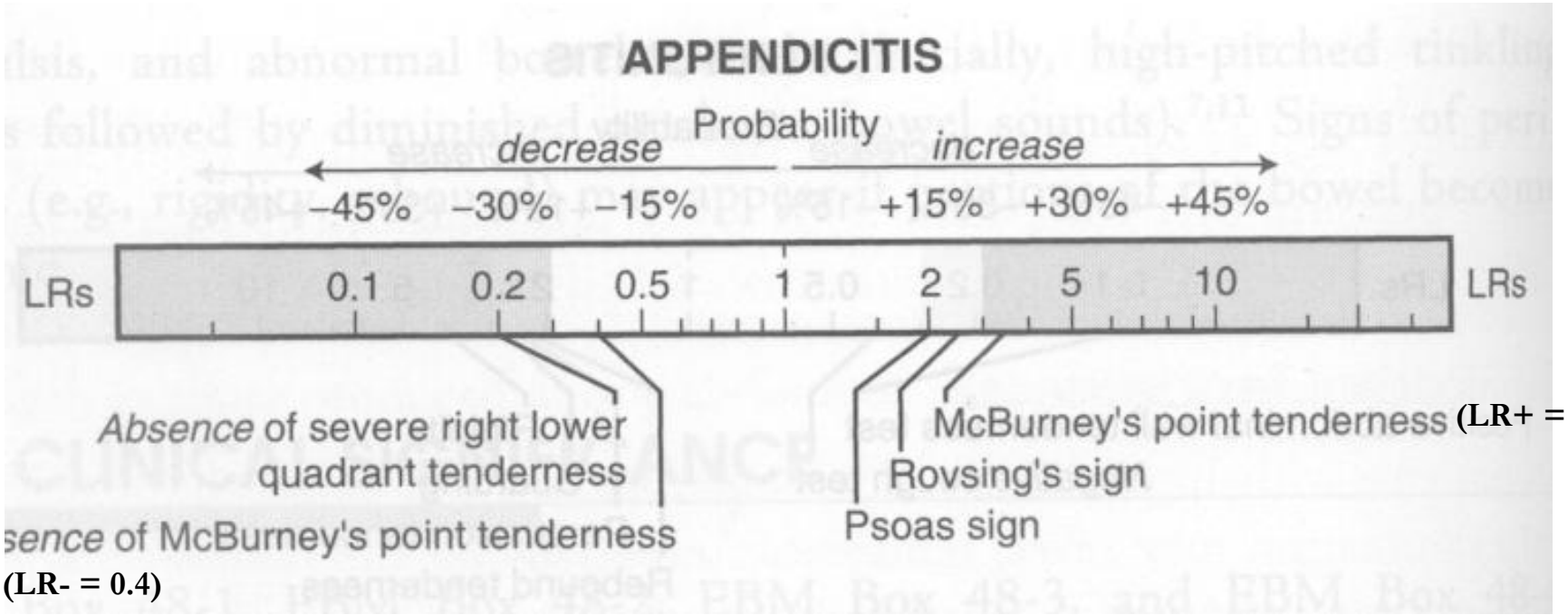
Rovsing's sign

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

For Example



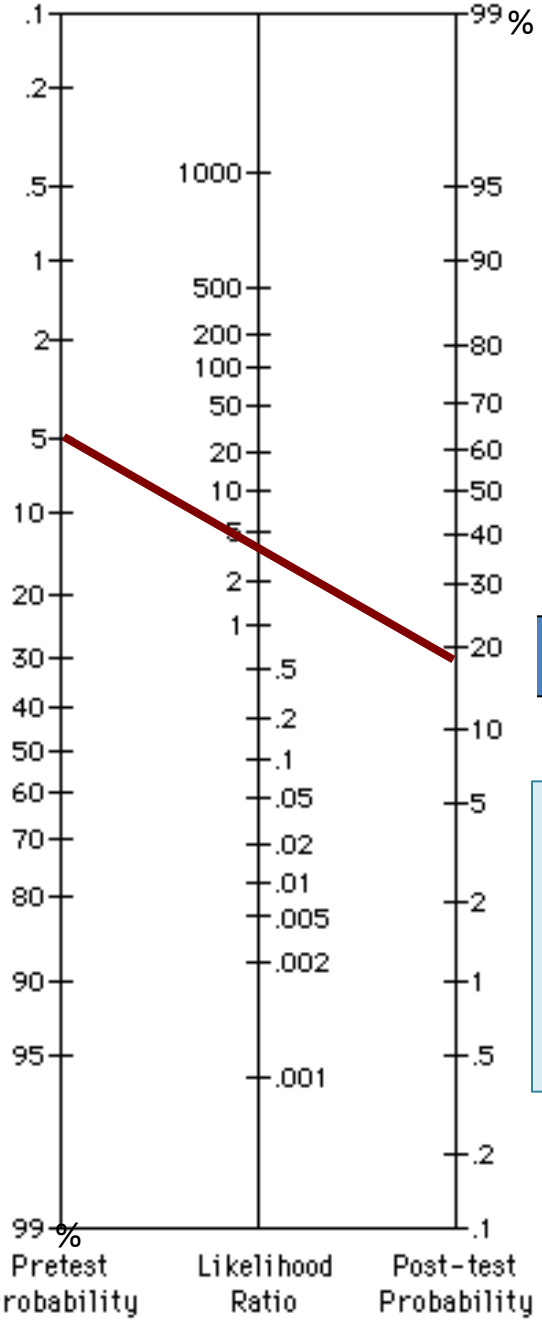
(LR+ = 3.4)

(LR- = 0.4)

McGee: Evidence based Physical Diagnosis (Saunders Elsevier)

Bayesian reasoning

Fagan nomogram



Pre test 5%

?Appendicitis:

McBurney tenderness LR+ = 3.4

Post test ~20%

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

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

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?

Will the test apply in my setting?

- Reproducibility of the test and 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?

BBC Mobile

News | Sport | Weather | iPlayer | TV

NEWS HEALTH

Home | World | UK | England | N. Ireland | Scotland | Wales | Business | Politics | Health | Education | Sci/Envir

10 August 2010 Last updated at 22:01

New brain scan to diagnose autism

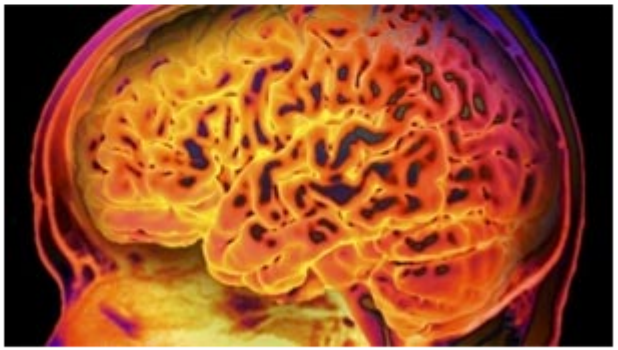
By Jane Hughes
Health correspondent, BBC News

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

The researchers detected autism with over 90% accuracy, the Journal of Neuroscience reports.

Natural Frequencies

Your patient asks you:

"If my child had this brain scan and it was positive, what's the chance my child has autism?? "

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. Murphy[1]

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

The National Autistic Society

Estimated prevalence rate in the UK

The indication from recent studies is that the figures cannot be precisely fixed, but it appears that a prevalence rate of around 1 in 100 is a best estimate a best estimate of the prevalence in children. No prevalence studies have ever been carried out on adults.

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

Morphometric feature	Correctly classified (%)	Sensitivity (%)	Specificity (%)	*p*
Left hemisphere				
All parameters	85	90	80	0*
Cortical thickness	90	90	90	0*
Radial curvature	72.5	65	80	<0.001
Average convexity	70	75	65	<0.004
Metric distortion	80	80	80	0*
Pial area	77.5	70	85	0*
Right hemisphere				
All parameters	65	60	70	<0.03
Cortical thickness	60	65	55	<0.01
Radial curvature	52.5	50	55	<0.30
Average convexity	50	40	60	<0.40
Metric distortion	57.5	45	70	<0.06
Pial area	45	45	45	<0.60

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.

UNIVERSITY OF OXFORD

CEBM CENTRE FOR EVIDENCE BASED MEDICINE

Natural Frequencies

Autism has a prevalence of 1%.

The test has sensitivity of 90% and specificity of 80%.

- 100% Always
- 50% Maybe
- 0% Never

Natural Frequencies

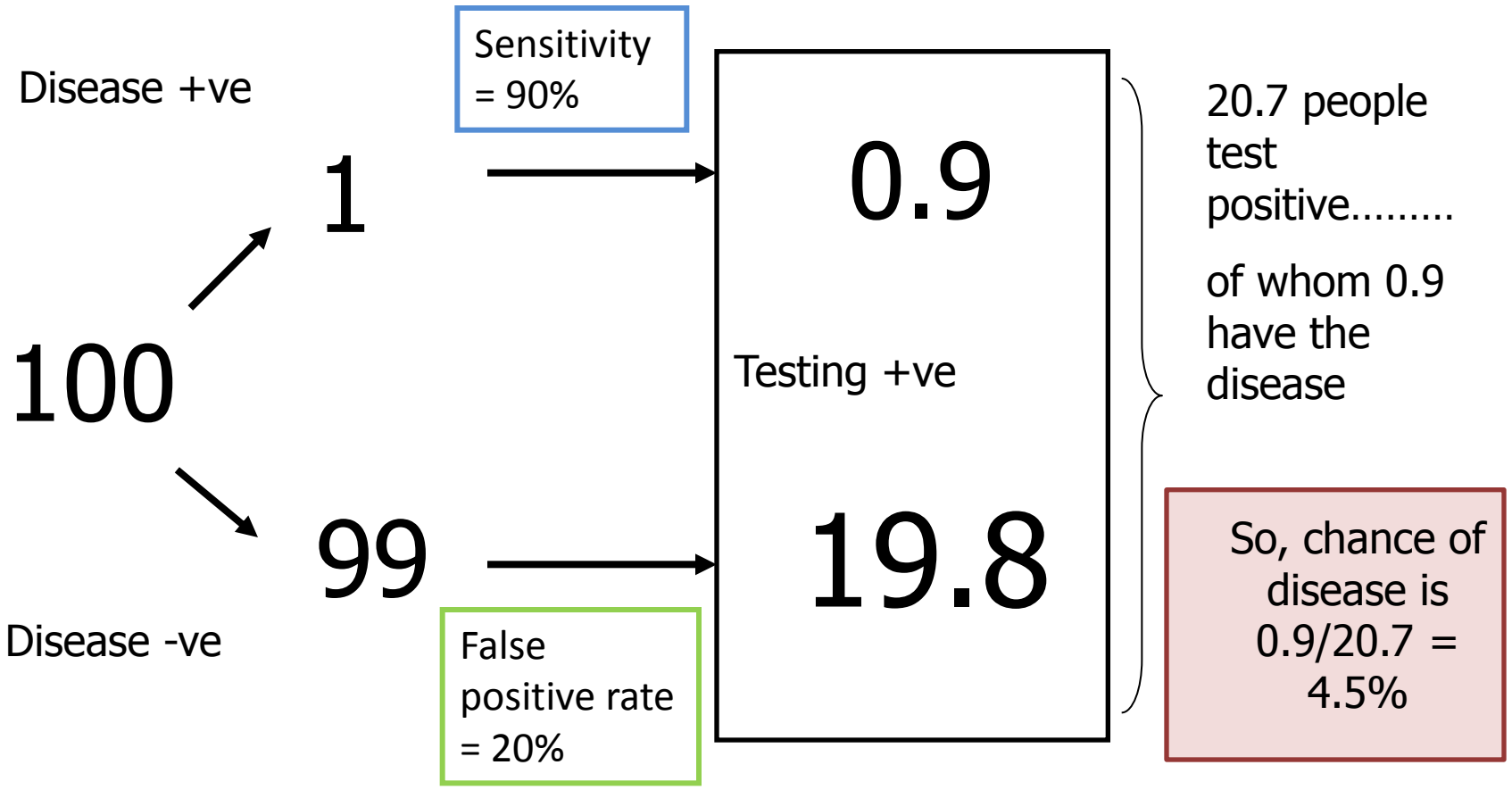
Autism has a prevalence of 1%.

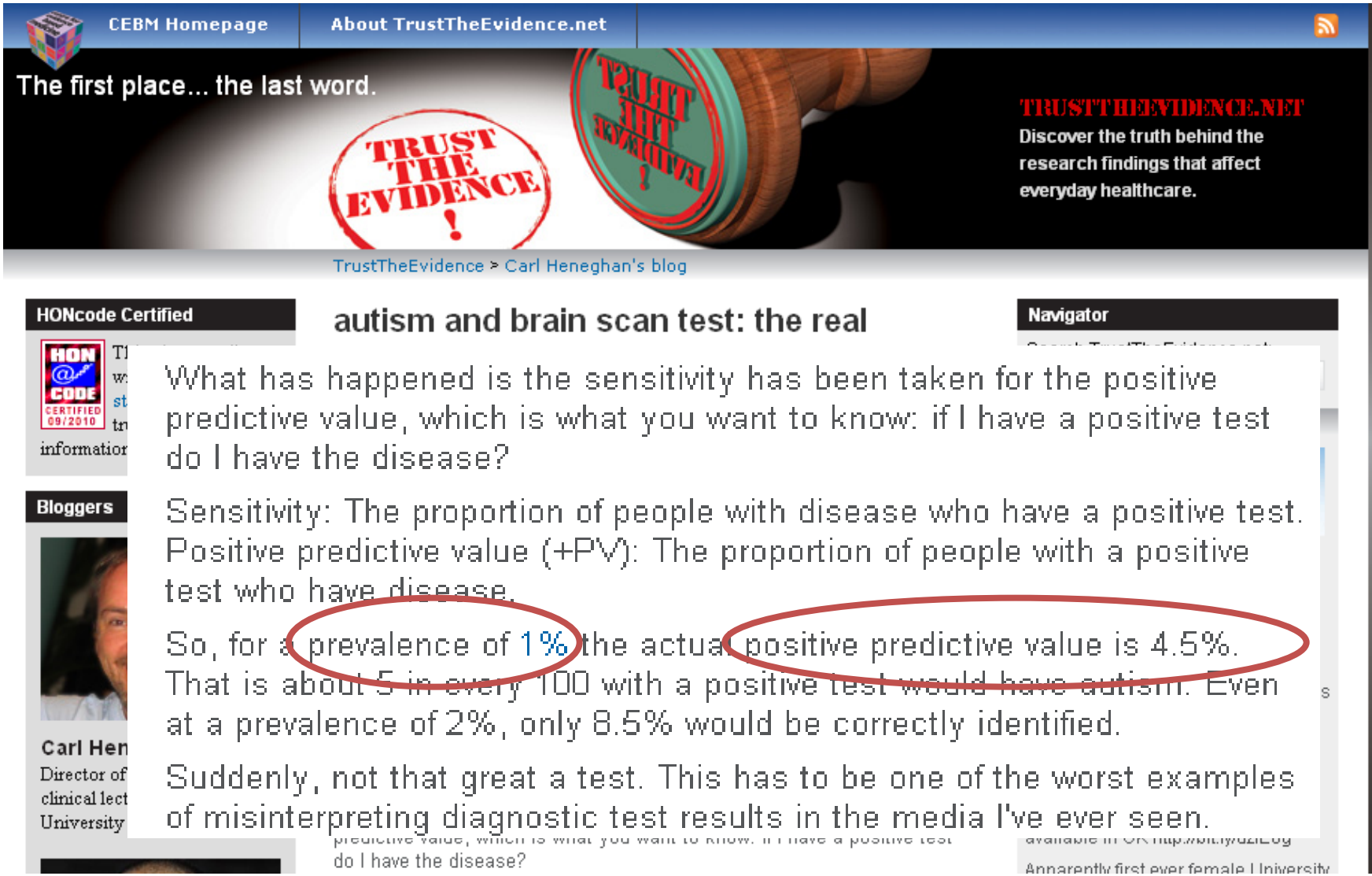
The test has sensitivity of 90% and specificity of 80%.

Given a positive test, what is the probability the child has autism?

End

Prevalence of 1%, Sensitivity of 90%, Specificity of 80%




CEBM Homepage
About TrustTheEvidence.net
The first place... the last word.
TRUST THE EVIDENCE !
TRUSTTHEEVIDENCE.NET
Discover the truth behind the research findings that affect everyday healthcare.
TrustTheEvidence > Carl Heneghan's blog
HONcode Certified
Navigator
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.
Bloggers
Carl Hen
Director of
clinical lect
University
do I have the disease?
Apparently first ever female University

NOTES&THEORIES
DISPATCHES FROM THE SCIENCE DESK

Previous Blog home Next

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.

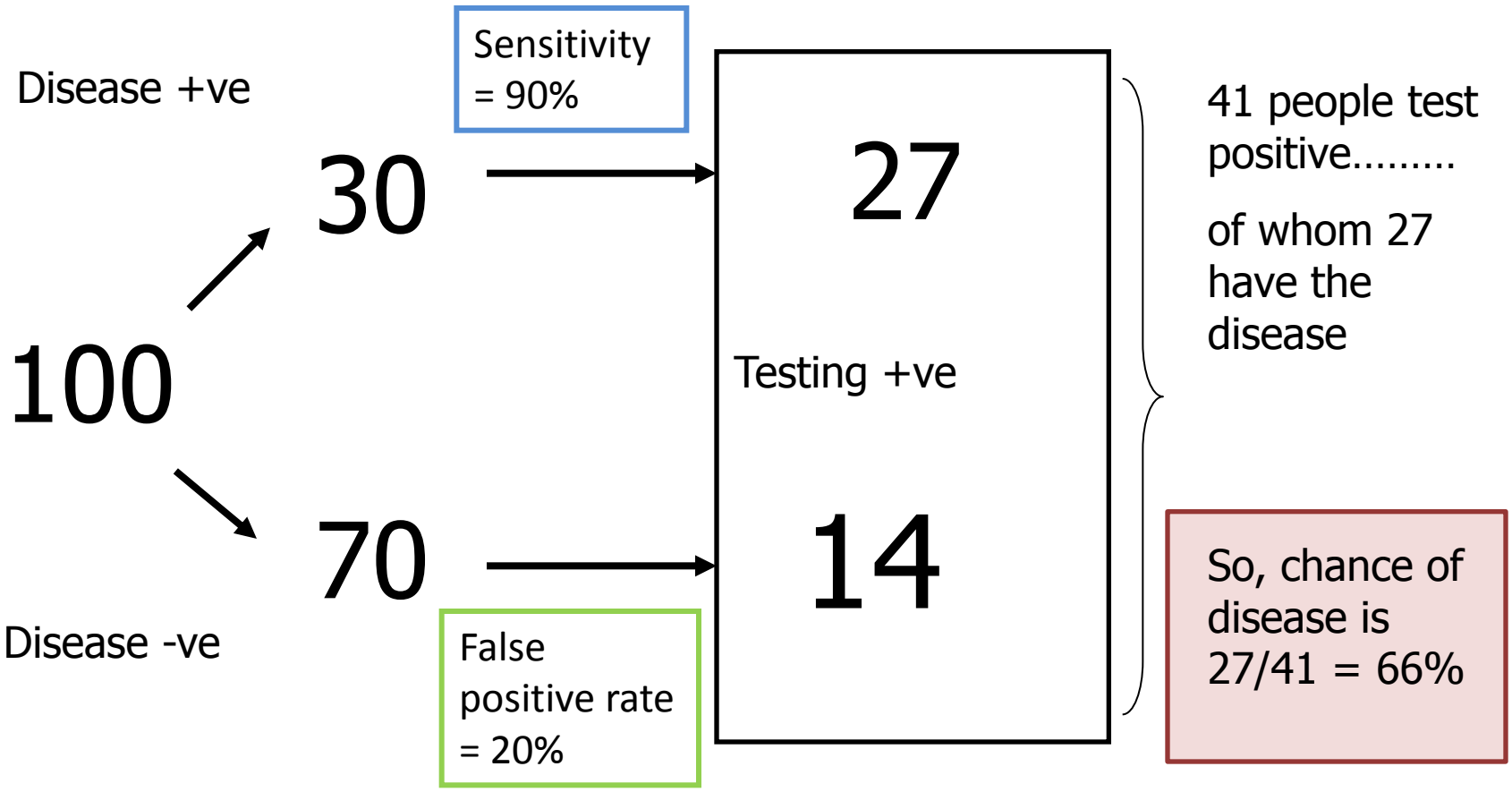
Well, you couldn't be any further from the truth.

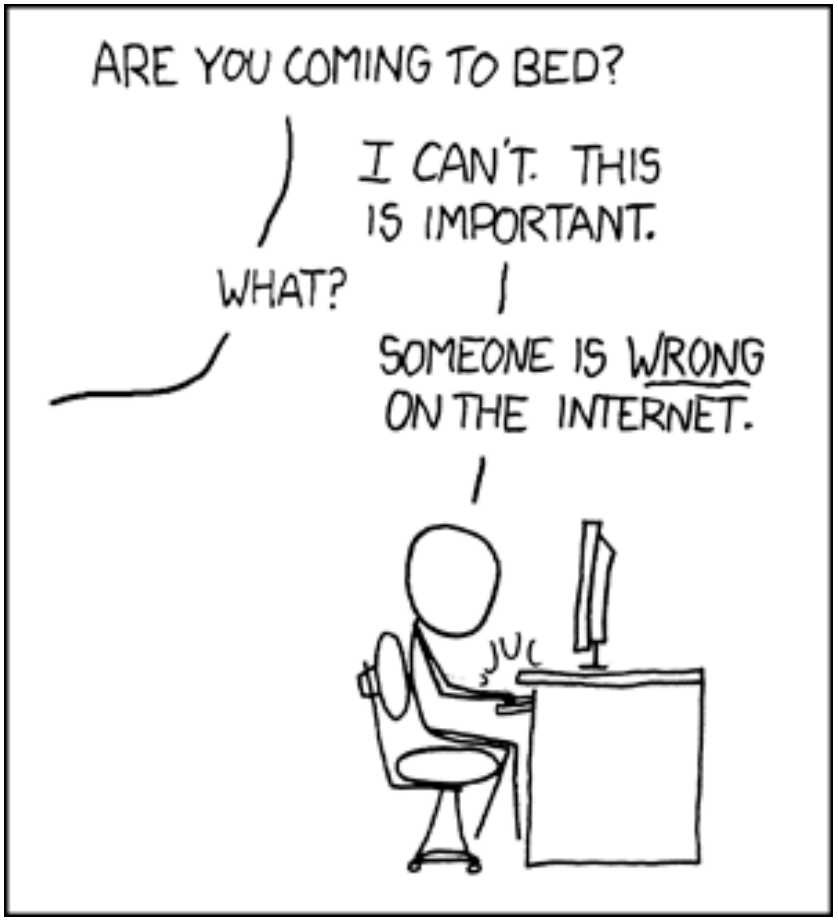
Posted by
Carl Heneghan Thursday
12 August 2010
15.29 BST
guardian.co.uk

Science
Medical research ·

Try it again….

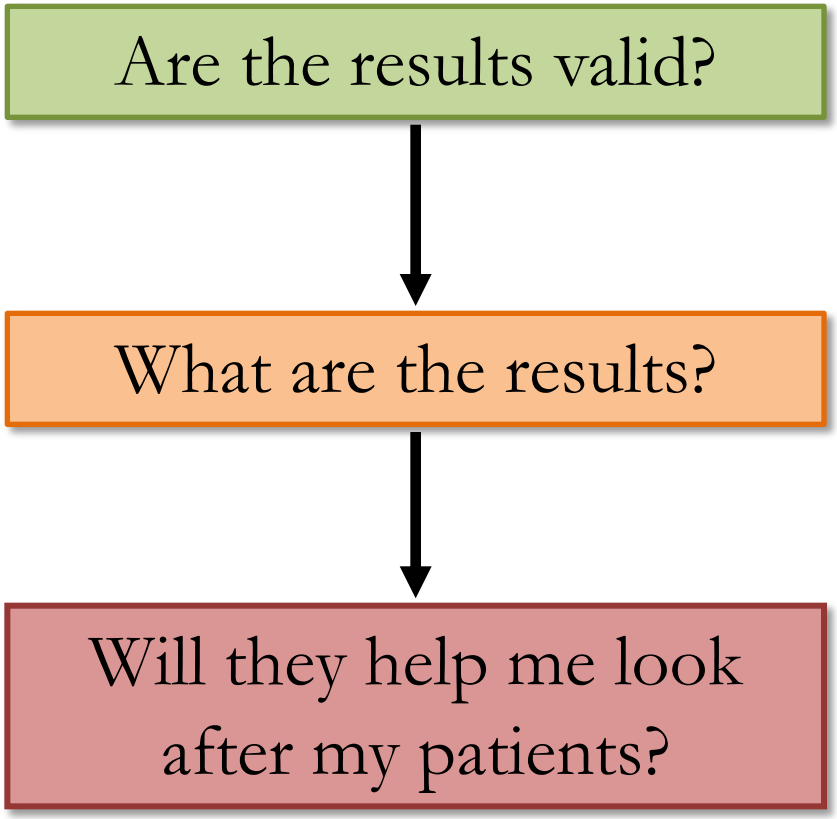
Prevalence of 30%, Sensitivity of 90%, Specificity of 80%



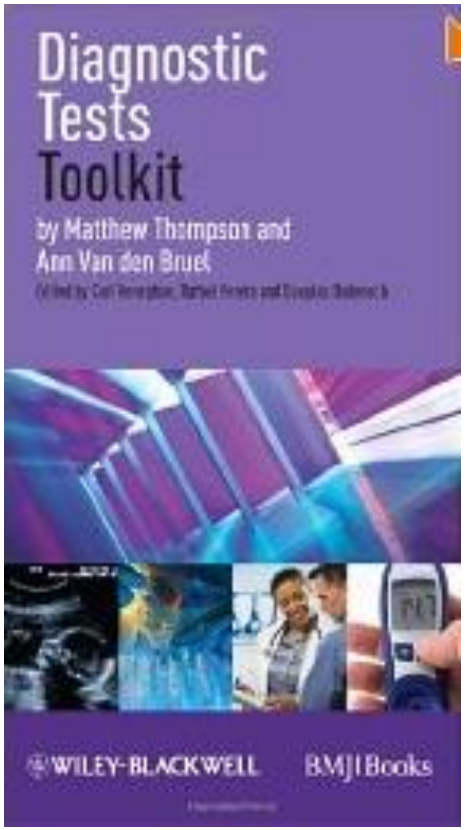

ARE YOU COMING TO BED?
I CAN'T. THIS IS IMPORTANT.
WHAT?
SOMEONE IS WRONG ON THE INTERNET.

www.xkcd.com

What is the ONE thing I need to remember from today?



Additional Resources



- Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. Brozek JL, Akl EA, Jaeschke R, Lang DM, Bossuyt P, Glasziou P, Helfand M, Ueffing E, Alonso-Coello P, Meerpohl J, Phillips B, Horvath AR, Bousquet J, Guyatt GH, Schünemann HJ; GRADE Working Group. Allergy. 2009;64(8):1109-16.

- QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. Ann Intern Med. 2011;155(8):529-36.
- Quality assessment tool for diagnostic accuracy studies: http://www.bris.ac.uk/quadas/quadas-2/

Now go and try it at home…..

…or in your small groups.