Welcome to the first workshop in a new course on Clinical Reasoning
READ CROSKERRY CH INTRO
What is ‘clinical reasoning’?

Clinical reasoning describes the thinking and decision-making processes associated with clinical practice.
What are the different elements of clinical reasoning? – mention proposed syllabus

This is my definition. It starts with clinical skills – how we communicate, get a history and examine the patient … and use all that information to estimate **how likely** a particular diagnosis is. This is known as the pre-test probability. And in general we are very bad at estimating probability, which is important because

The next element of clinical reasoning is the correct use and interpretation of diagnostic tests. But tests give us test probabilities, not real probabilities. Tests rarely tell us what is wrong with a patient – they are not a ‘yes’ or a ‘no’.

Another key element is understanding cognitive biases and human factors. Human thinking is flawed and clinicians are no exception.

Which leads us on to another element – the ability to think about one’s own thinking, to know what mode of thinking you are engaged in, different factors that affect our thinking, and how we might act to mitigate these.

Clinical reasoning does not end with a diagnosis. Patient-centred evidence based medicine is also an element of clinical reasoning – that means applying the available evidence to the patient in front of you.

**Clearly knowledge is required.**
Clinical reasoning assessments

- MCQs
- Formative CR assessments in the 1st year
- A clinical scenario based on content you have covered
- Unlike other assessments, you get more marks for showing what you are thinking and why rather than the ‘right answer’
Resources: there’s a website with recommended books/resources and lecture notes (as well as on Moodle).
You can follow the clinical reasoning blog on Twitter ...
Why is clinical reasoning important?

- Diagnosis is wrong 10-15% of the time
- Diagnostic error is more likely to lead to harm than other types of error
- Two thirds of the root causes of diagnostic error involve errors in reasoning – most commonly when the available data is not synthesised correctly

Why is clinical reasoning important?
As we have seen there is more to CR than diagnosis but if we just focus on the diagnostic process for a moment ...
It may surprise you to hear that diagnosis is wrong 10-15% of the time. 85% of diagnostic errors are considered preventable. Serious disability results in up to 50% of cases of diagnostic error.
For every diagnostic error there are a number of root causes, but errors in reasoning account for 2/3 of these.
Studies of diagnostic error assign three main categories: no fault errors, system failures, and human cognitive error. But if you look more closely at “human cognitive error” what do we find?

1] The first is knowledge gaps. The best way to increase your knowledge and diagnostic skill through as much experience as possible, deliberate practice and feedback. You cannot learn medicine from books and tutorials, you have to build up a vast databank of clinical cases (virtual and real) where you get the history, examine the patient, look at initial test results and practice NAMING the problem(s) and what you are going to do about it.

2] The second is misinterpretation of diagnostic tests – the topic we are going to cover today.

3] The third is cognitive errors and biases – topics we will cover in future workshops.
So without further ado, let’s take a closer look at ONE element of clinical reasoning: use and interpretation of diagnostic tests.

Some tests we use in medicine are fairly straightforward: FBC

Some are more complicated and need to be interpreted – as you will see!
Clinical skills are really important

Studies over the last 30 years show that history and examination make the diagnosis 70-90% of the time, even with the explosion of medical technology that has taken place.

The first really important concept for you to grasp is that history (and examination) is everything! And don’t forget history is not only from the patient – eye-witnesses, carers and relatives are also sources of history. While you are taking a history and examining a patient you have to be able to INTERPRET the information you are getting and be good at getting information ... you have to have good clinical skills.

The second concept for you to grasp is that you cannot reason without medical knowledge. However, there comes a point when you have enough medical knowledge but if you cannot use that knowledge, think logically and reason effectively, you won’t be a good doctor. This is what this course is about – how to use that knowledge.

The acquisition of knowledge and reasoning skills go hand in hand and TAKE TIME. That’s why we have a website so you can go back to things throughout your medical course as your knowledge and experience grows.
A really important concept for you to grasp today is that there is no such thing as a perfect test – test results alone are not the answer! Think about this a bit more when you are in your PBL groups.

[read slide]

We are going to explore each of these points in a bit more detail and get you to work on a problem in small groups later ...
We define what a normal test result is based on studies of the population. **Normal is arbitrary.**

If someone is just outside the normal range they may not have a disease, and if someone is in the normal range they could still have a disease. Let’s take ferritin as an example – we measure that to look for iron deficiency. First, values for men and women are different.

Second, it’s considered normal (not a disease) for young women to have a degree of iron deficiency (because of menstruation). On the other hand, ferritin is also an ‘acute phase protein’ – it goes up in infection or inflammation, so if you are looking for iron deficiency in someone who happens to have a pneumonia, the result may well be falsely normal.

In a severe asthma attack a low PaCO2 is to be expected (due to hyperventilation). A “normal” PaCO2 in a severe asthma attack is very bad ... (do you understand why?)
Factors other than disease that influence test results

- Age
- Ethnicity
- Sex
- Pregnancy
- Body position
- Spurious (in vitro) results

Continuing on that note, there are factors other than the disease you are looking for that influence test results.
Read slide / elaborate.

For example, alk phos is raised in teenagers, 2 year olds have low Hbs, elderly creatinine can be normal with reduced eGFR
ECGs of young black males are different, pregnant women are anaemic etc.
Operating characteristics

• Patients need to be able to comply with some tests
  – Spirometry
  – Exercise tests
• Some tests are very ‘operator dependent’ or influenced by body habitus
  – Echocardiography
  – Ultrasound
• Some conditions are paroxysmal
  – Epilepsy (EEG)
  – Asthma

Tests are also subject to what’s called ‘operating characteristics’
Read slide / elaborate
All tests have a characteristic called sensitivity and specificity.

There is no test that can 100% of the time detect people with a disease and exclude people without it.

Even the best tests with say 90% sensitivity and specificity will miss 10% of people with the disease and incorrectly identify 10% of normal people as having the disease.

**So we have this concept of ‘true positives’ and ‘false positives’.** That’s really important to understand.

A very sensitive test will detect most disease but may generate abnormal findings in healthy people; a negative result will therefore reliably exclude disease, but a positive test is likely to require further evaluation. A very specific test may miss significant pathology but is likely to establish the diagnosis beyond doubt when the result is positive.

Now, all tests are different – so as a medical student and a doctor, you should try to at least have a rough idea about how good a test is for the thing you are looking for. It might be good, it might be terrible!
The probability of a disease depends on the clinical (pre-test) probability plus the sensitivity and specificity of the test.

An elderly lady is admitted following a fall. She had hurt her left hip and was unable to weight bear. On examination, the left hip was extremely painful to move. Her X-ray (shown) is normal. Is there a fracture?

Here’s a simple clinical example ... (vote)
A shortened, externally rotated leg, severe pain (unable to weight bear) are hallmarks of fractured neck of femur.

YES, there is a fracture! In a patient with a high clinical probability a normal result does not exclude the disease. However, in a patient with a low clinical probability, a normal result CAN exclude the disease – in this case a fracture.

So important principle: the probability of a disease depends on the clinical (pre-test) probability ... plus the sensitivity and specificity of the test. XRs cannot always detect fractures.

In this case, repeat AP + lateral X-rays or a CT or MRI will show up the problem.
The most fundamental principle in clinical decision making is that the interpretation of new information depends on what you believed beforehand.

Cox and colleagues state that the most fundamental principle in clinical decision making is that the interpretation of new information depends on what you believed beforehand. In other words, the interpretation of a test result depends on the probability of the disease before the test. Clearly you need good medical knowledge of how diseases present plus their epidemiology to estimate the pre-test probability.

They go as far to say, ‘Once you accept this principle, your life will never be the same again.’

Tests are commonly misused by clinicians. We do not understand probabilities or the information we receive from tests. Tests change the probability of a particular disease being present or absent, but rarely in a binary yes/no fashion. More commonly a test will increase or decrease the likelihood of a disease being present by less than we think.

Use stroke and CT head as another example

Now while I don’t want you to go around disbelieving tests (!) the point here is that tests have to be INTERPRETED. Tests don’t make a diagnosis, doctors do.
Does this patient have emphysema?

- You see a 60 year old ex-miner who has smoked 20 cigarettes a day for the last 40 years.
- He has persistent breathlessness on climbing hills and stairs, with wheezing.
- You send him for a spirometry test, commonly used in this situation. The test result is normal.
Probability of having a disease

http://vassarstats.net/clin2.html
Probability of having a disease

http://vassarstats.net/clin2.html
Any questions so far?
Prevalence of disease

If a test to detect a disease whose prevalence is 1:1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person’s symptoms and signs?

(45% of Harvard doctors said 95%)

Now let’s get more complicated.
Have a think about this problem ...
As you can see, almost half of very smart doctors said 95% - but they are only thinking about sensitivity and have neglected to take prevalence in to account.
Answer

False positives 50/1000
But only 1/1000 actually has the disease
Means the chance of a positive result with disease = 1 out of 51 or 2%

<table>
<thead>
<tr>
<th></th>
<th>Present</th>
<th>Absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual disease</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td>Test +</td>
<td>1</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>Test -</td>
<td>0</td>
<td>949</td>
<td>949</td>
</tr>
</tbody>
</table>

The importance of understanding prevalence (or denominator).

Read slide.
Read slide.

Positive and negative PVs are influenced by the prevalence of the disease in the population being tested.

If a patient walks in to ED with a fever and aches and pains, it could be anything – but if a patient walks in to ED with the same symptoms during a flu pandemic, he is highly likely to have flu.
Example of an imperfect test for something that is not very common ...

A cab was involved in a hit-and-run at night. Two cab companies operate in the city, the Green and the Blue. 85% of the cabs in the city are Green and 15% are Blue. A witness identified the cab as Blue. The Court tested the witness under the circumstances that existed on the night of the accident and concluded that the witness correctly identified the colour 80% of the time. What is the probability that the cab was actually Blue?

(The most common answer is 80%)

So let’s get you to combine the principles we have just covered and solve some problems. The first is a non-clinical example I’ll talk you through, and the second is a clinical example you will work out in small groups.

This problem is from the book, “Thinking, fast and slow” which is basically about how illogical humans can be.

There are two items of information here: a base rate and the imperfect reliability of the witness (the test).
If there was no witness, the probability of the guilty cab being blue is 15%.
If the two cab companies were equally large, then the base rate is irrelevant and we only need to consider the reliability of the witness, so the answer is 80%.
But we have to combine these two sources of information.

Statistical base rates are facts about a population to which a case belongs, but they are not relevant to the individual case. They are generally underweighted and sometimes neglected altogether when specific information about the case is available.
For example – temporal arteritis is a condition that only exists in people aged over 55. A 25-year-old woman came with a headache just like temporal arteritis and even had a blood test result consistent with this diagnosis … what did the house officer diagnose? (temporal arteritis – ignoring the fact that this was impossible in this particular patient).
Answer

<table>
<thead>
<tr>
<th></th>
<th>Green</th>
<th>Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>85</td>
<td>15</td>
</tr>
<tr>
<td>Witness +</td>
<td>68 (80% of 85)</td>
<td>12 (80% of 15)</td>
</tr>
<tr>
<td>Witness -</td>
<td>17</td>
<td>3</td>
</tr>
</tbody>
</table>

- 12/100 times the witness will correctly identify a Blue cab as Blue
- 17/100 times the witness will incorrectly identify a Green cab as Blue
- There is therefore a 12+17=29% chance the witness will identify the cab as Blue
- This results in a 12/29 or 41% chance that the cab identified as Blue is actually Blue.

Base rate neglect (prevalence neglect)

Talk through slide ...
Small group work

You see two patients with chest pain and decide to send them both for an imaging stress test to see whether they have angina (see WORKSHEET). Stress testing has a sensitivity of 90% and specificity of 85%

We know the actual prevalence of IHD in the population based on angiography and PM studies. What is the chance of a positive stress test being correct in each of your two patients? (The answer is not 90%).

Acknowledgement.

Stress testing has a sensitivity of 90% and specificity of 85%. That’s quite typical of a diagnostic test.

The first patient is a 65 year old man with a history of typical angina. The second is a 35 year old woman with atypical (non-anginal) pain. First of all the clinical probability is different in these two patients – one has a good history, the other sounds like it’s not really angina. Secondly, the prevalence of IHD in 65 year old men is very different to the prevalence of IHD in 35 year old women (as illustrated by the slide). So we need to consider all these factors when deciding how to interpret their test result.

<table>
<thead>
<tr>
<th>Age</th>
<th>Non-anginal CP Men</th>
<th>Non-anginal CP Women</th>
<th>Atypical angina Men</th>
<th>Atypical angina Women</th>
<th>Typical angina Men</th>
<th>Typical angina Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-39</td>
<td>5.2</td>
<td>0.8</td>
<td>21.6</td>
<td>4.2</td>
<td>68.7</td>
<td>25.8</td>
</tr>
<tr>
<td>40-49</td>
<td>14.1</td>
<td>2.8</td>
<td>46.1</td>
<td>13.3</td>
<td>87.3</td>
<td>55.2</td>
</tr>
<tr>
<td>50-59</td>
<td>21.4</td>
<td>8.4</td>
<td>58.9</td>
<td>32.4</td>
<td>92</td>
<td>79.4</td>
</tr>
<tr>
<td>60-69</td>
<td>28.1</td>
<td>18.6</td>
<td>67.1</td>
<td>54.4</td>
<td>94.3</td>
<td>90.6</td>
</tr>
</tbody>
</table>
In a patient with such a high clinical probability, a negative test result is almost twice as likely to be wrong than right!

<table>
<thead>
<tr>
<th></th>
<th>IHD</th>
<th>No IHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>94</td>
<td>6</td>
</tr>
<tr>
<td>Pos ST</td>
<td>84.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Neg ST</td>
<td>9.4</td>
<td>5.1</td>
</tr>
</tbody>
</table>

In a patient with a high clinical probability, you can trust the results 90% of the time a positive stress test is correct. A negative result should be considered with suspicion.
In a patient with such a low clinical probability, you cannot necessarily trust the results.
There are far more false positive stress test results than true positives - makes you wonder, doesn’t it! A positive result does NOT mean this person has angina/IHD. If a test does not have 100% specificity – don’t do it on a patient with such a low clinical probability!
Now, IMPORTANT – not all tests perform in the same way. Some are good and some are bad at what you are looking for. Some results are more trustworthy than others – for example, you can nearly always rely on a full blood count result (within a small margin of error). But more complicated tests (X-rays/ultrasound/lung function etc) require interpretation!

### 35 year old woman with atypical chest pain: stress test results

<table>
<thead>
<tr>
<th></th>
<th>IHD</th>
<th>No IHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>1</td>
<td>99</td>
</tr>
<tr>
<td>Pos ST</td>
<td>0.9</td>
<td>14.9</td>
</tr>
<tr>
<td>Neg ST</td>
<td>0.1</td>
<td>84.1</td>
</tr>
</tbody>
</table>

In a patient with such a low clinical probability, a positive test result is 15 times more likely to be wrong than right!
A test result by itself is not the answer
(unfortunately patients don’t know this!)

- Tests must be interpreted in the light of CLINICAL PROBABILITY
- You must also know something about the CHARACTERISTICS of the test in question
- And if the PREVALENCE of the disease is very high or very low in the patient’s group – this affects the predictive value of the test
What happens when the clinical probability and the test results don’t match?

- Test results on patients with extremely low or high prior probabilities can be misleading
- When a test result confirms a clinical suspicion then it is likely to be true
- When a test result contradicts a clinical suspicion, more thinking may be required ...

Read slide

**Which is why it’s really important not to use tests indiscriminately or unwisely**
Therapeutic threshold

• The therapeutic threshold combines factors such as test characteristics, risks of the test, and the risks and benefits of treatment.
• The point at which all factors are evenly weighed is the threshold.
• If a test or treatment is effective and low risk we would have a lower threshold for going ahead ...

Last concept: the *therapeutic threshold* combines factors such as test characteristics, risks of the test, and the risks vs benefits of treatment. The point at which the factors are all evenly weighed is the threshold. If a test or treatment for a disease is effective and low risk then one would have a lower threshold for going ahead. On the other hand, if a test or treatment is less effective or high risk, one requires greater confidence in the clinical diagnosis and potential benefits of treatment first.

  e.g. urinalysis in patient’s with clinical pyelonephritis – urinalysis (dipstick) has 75-85% sensitivity in this condition.

  So in clinical pyelonephritis I give antibiotics while I gather more information e.g. with an MSU, renal tract ultrasound, maybe blood cultures …

  But if I was faced with a possible lung cancer I would require a biopsy and staging CT before starting treatment …
SUMMARY

Test results can be affected by things other than disease (such as age, ethnicity and operating characteristics). Assessing clinical (pre-test) probability is vital. Without it, you cannot interpret any test result.

The predictive value of a test result not only depends on the test’s sensitivity and specificity, but also on the prevalence of the disease in the population studied.
Discussion

www.clinical-reasoning.org