Clinical investigation – the theory

This study programme is about medical imaging, so we'll be talking about imaging investigations. However, just about everything we say about imaging will apply to any other form of clinical investigation – blood tests for example.

Patient perceptions

Patients are used to the concept of having 'tests' carried out, and have great faith in the power of clinical investigations to solve the problem that took them to the doctor. However, this faith is often misplaced.

Most people (and not just patients!) make a number of assumptions about clinical tests:

- there are only two outcomes to a test: it's either positive or negative; either they have got the disease the doctor suspects, or they haven't
- if the disease is present, the test will reveal it
- if they haven't got the disease, the test will come up negative

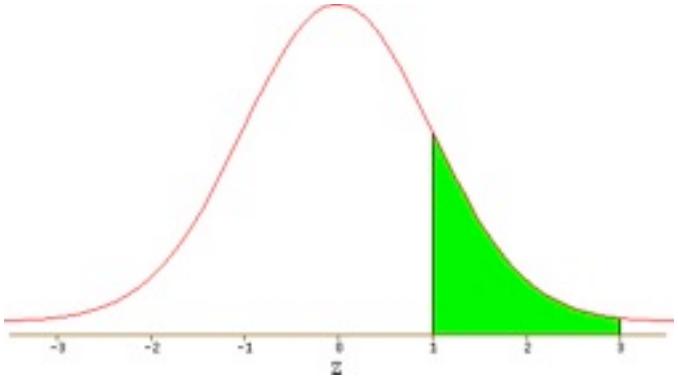
There are a number of problems with these assumptions, not least of which is the fact that none of them are necessarily true. We can discover nearly everything we need to know concerning the theory of clinical investigation by considering each of the assumptions in turn.

1. positive or negative? – a simple choice

Well, in some cases, yes. For example, a blood test for HIV will usually give a fairly unequivocal result; you are either +ve or -ve (of course, the result may be wrong, but we'll come on to that later). Some imaging results are pretty clear-cut as well – a straightforward fracture, for example. But even there, we may end up saying we're not sure whether there is a fracture present or not.

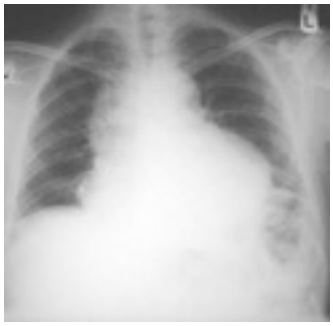
example 1: Take a simple blood test. We all worry about our blood cholesterol levels, because we know that too high a level can be associated

with an increased risk of what the cardiologists euphemistically call 'adverse cardiac events' (you know, death, that sort of thing). It is even possible to buy home testing kits to frighten yourself with. This all assumes that it is easy to identify a threshold level to separate normal from high values (below X mmol/l you can stop worrying, greater than X, and you have to start dieting or taking cholesterol-lowering drugs). But if you look at the distribution of cholesterol levels in the population, the result would be something like the normal distributionin the following graph:



So where do you draw the line separating normal and high (green) levels? Too low, and a lot of people who will never run into problems due to their cholesterol will be treated (and worried) unnecessarily and at great expense. Too high, and people who would have benefited from cholesterol-lowering treatment will miss out.

<u>example 2:</u> In many radiological situations too, the key features of the image we see will be situated somewhere on a spectrum between definitely normal and definitely abnormal, and reporting becomes more of an art than a science. This chest X-ray for instance:



The heart is enlarged. Is it worth mentioning in the report? Well, it depends on a number of things, not least the age of the patient and the symptoms they are complaining of. If it's an eighty year old with no symptoms related to the heart, it will pass for normal. If it's a forty year old man with sudden onset of chest pain, we would take it more seriously.

So as far as positive/negative is concerned, we are talking shades of grey, not black and white, so assumption 1 is false for most real clinical situations.

2. if disease is present, the test will pick it up

Maybe, maybe not. Here we are talking about the sensitivity of the test. If you take 100 people who all have a disease, and the test you are using turns up positive in all of them, it has a sensitivity of 100%. If it is negative in 10 of them, despite the fact that they have the disease, the sensitivity of the test is 90/100 = 90%.

Very few tests have a sensitivity of 100%, so assumption 2 is false.

3. if the disease is absent, the test will be negative

Again, maybe. This time we are talking about a quantity called specificity. Let's take 100 patients who are disease-free, and test them. If the test is negative in all 100, the specificity is 100%. If it is positive in 10 of them, despite the fact that they don't have the disease, the specificity of the test is 90/100 = 90%.

Very few tests have a specificity of 100%, so assumption 3 is false as well.

This illustrates an important truth which underpins the way we use imaging in diagnosis:

there is no such thing as a perfect test

In the next section, we look in more detail at the calculation of sensitivity and specificity and the relationship between them, and also at a couple of other parameters of test performance which are more helpful to us when we use medical imaging in clinical practice.

sensitivity and specificity

When talking about the efficiency or accuracy of tests, we have to start by looking at the possible results of the test in patients who do, or do not, have the condition you are trying to diagnose. Those results break down as follows:

test is positive; patient has disease: this is a true positive (TP) result
test positive; patient disease-free: this is a false positive (FP) result

test is negative; patient disease-free: this is a true negative (TN) result
test is negative; patient has disease: this is a false negative (FN) result

obviously: TP & TN results are good because they are accurate
FP & FN results are bad because they are misleading

sensitivity is the proportion of those with the disease who test positive:

$$\frac{TP}{TP + FN}$$

specificity is the proportion of those without the disease who test negative:

$$\frac{TN}{TN + FP}$$

So: sensitivity tells you how good a test is at detecting disease
specificity tells you how good it is at excluding it

or: sensitivity tells you how many false negatives the test will produce
(80% sensitivity means 20% false negative results)
specificity tells you how many false positives to expect
(85% specificity means 15% false positive results)

There is an important inverse relationship between sensitivity and specificity: i.e. as one goes up, the other goes down. This makes sense. Take the following example (a silly one, but it makes the point):

example: Suppose we are wanting to diagnose the condition of gigantism. This is caused by an excess of growth hormone during childhood, and results, not surprisingly in increased growth rates. Now suppose we choose as our diagnostic test the simple measurement of height at the age of fourteen. First we have to choose a cut-off point to separate out the children with the condition from their shorter normal friends, remembering that there will be a grey area which includes the taller normals and shorter giants.
if we choose 5ft 10in as the upper limit of normal: we will detect all the children with gigantism, because they will all be taller than this – there will be no false negatives and the sensitivity will be 100% (good). However, some of the taller normal kids will also be in this height range, so there will be quite a few false positives, lets say 8% - the specificity will therefore be 92% (not so good).
if we raise the cut-off to 6ft 2in: a number of the affected children will be shorter than this, let's say 20%, so the sensitivity will be down from 100 to 80%. However, with the cut-off set so high, there will be no normal kids in the abnormal range, so no false positives, and a specificity of 100%.

So, there is a trade-off between sensitivity and specificity. This is important when deciding what test to use in any particular clinical situation, and we will see a good illustration of this when we consider screening tests, later.

predictive value

Sensitivity and specificity are all very well, but in real clinical situations they are often not very helpful. For example, if a doctor has sent a patient for a test, and the result has come back as positive, what he or she wants to know when the patient returns is not how sensitive or specific the test is, but how likely it is that the positive result is a true positive, and the patient really is suffering from the disease. Similarly, if the result was negative, the doctor wants to know how confident he can be in giving the patient the all-clear.

This is where predictive value comes in. The positive predictive value (PPV) is the likelihood that a patient with a positive test result actually does have the disease in question:

$$PPV = \frac{TP}{TP + FP}$$

The negative predictive value (NPV) is the likelihood that a patient with a negative test result really is disease-free:

$$NPV = \frac{TN}{TN + FN}$$

Predictive value is therefore very useful. The drawback is that it isn't constant for a particular test: it varies with the prevalence of the disease in the population from which the patient comes. The mathematical reasoning behind this is explained at the following website:

http://www.ipathology.com/ipathology/effect-of-disease-prevale.html

In other words :

- the lower the prevalence, the higher the predictive value of a negative result and the lower is the predictive value of a positive
- the higher the prevalence, the higher the predictive value of a positive result and the lower the predictive value of a negative

The relationship between the prior probability of disease, the test result and the resulting post-test probability that the patient has the condition in question has been formalised as Bayes' Theorem. To understand this, you need also to understand the concepts of joint and conditional probability, which sounds pretty complex, but in fact only utilises the simple concepts covered earlier in this section. One of the best illustrations of the application of Bayes' theorem to a real-life situation, along with a brief discussion of the principles involved, will be found at the following website: read it now.

http://www.maa.org/devlin/devlin_2_00.html

Putting the theory into practice

Bayes' theory is all very well, but if you write it down, it looks like this:

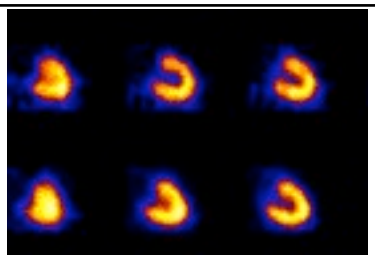
$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F')P(F')}$$

So, we don't sit in clinics with a calculator working these sums out. However, the formula above is only the mathematical expression of some general principles about clinical testing that we covered in the previous section. To apply Bayesian analysis to a clinical problem, we need to know the prior probability that a particular patient has the disease in question, and have an understanding of the way in which the predictive value of a test changes with that prior probability.

This is where we start to translate the fairly dry theory which has gone before into clinical practice with real patients, so read on.

We'll use a cardiology example to illustrate the theoretical principles around predictive value and decision making in clinical testing. You will take the role of the cardiologist about to review an outpatient who is suffering from chest pain which may be due to cardiac ischaemia (inadequate blood flow through the coronary arteries which supply the muscle of the heart wall).

You have taken a careful clinical history and examined him, he has had an electrocardiogram recorded (ecg: a tracing of the electrical activity of the heart which can provide evidence of ischaemia) and measurements of cardiac enzyme levels in his blood (these enzymes are released into the bloodstream when heart muscle becomes ischaemic). You will be reviewing the evidence at your disposal, and deciding whether the patient needs another test to confirm or refute the diagnosis of ischaemic heart disease. That test is a nuclear medicine perfusion scan:



cardiac perfusion scan, with normal isotope distribution throughout the left ventricle, indicating normal blood flow.

We will consider two patients, Mr A and Mr B, who are very different (see boxes 1a and

1b, below:

box 1a

Mr A

- 30 years old, generally fit, slim
- keen sportsman
- complaining of chest pain which could be due to his heart, but which is not typical
- no family history of heart disease
- non-smoker
- ecg: not normal, but again not typical of ischaemic change
- cardiac enzyme levels normal

box 1b

Mr B

- 55 years old, overweight
- bus driver, takes no exercise
- complaining of typical anginal (cardiac) chest pain
- all his immediate male relatives died by age of 65 from heart disease
- heavy smoker
- ecg: typical changes of ischaemia
- cardiac enzyme levels raised

task 1: answer the following questions, first in relation to Mr A, then to Mr B:

a) what is the prior probability (low, intermediate or high) that the patient's symptoms are due to ischaemic heart disease?

b) using the same rating scale, what will be the predictive value of a negative result from the perfusion scan on each patient?

c) and will what will be the predictive value of a positive result?

Task 1 answers:

a) For patient A, the probability is low. He is young, fit, and has none of the risk factors for developing heart disease in early life. Also, the other tests he has had (ECG, enzymes) are normal. Mr B, on the other hand, has typical symptoms, all the risk factors, plus supportive evidence of heart disease from the other tests. The pre-test probability that he is suffering from cardiac ischaemia must be close to 100%.

b) For patient A, the NPV will be high, because the prior probability is so low. In other words, the scan result is in accordance with your pre-scan

impression that the patient doesn't have heart disease, and so you know it is likely to represent a true negative result.

For patient B, the reverse applies. Given his history and examination findings, the prior probability that he has heart disease is so high that the test result is more likely than not to represent a false negative. Another way of saying this is that the NPV is low.

c) For patient A, the PPV would be low. In other words, the prior probability of disease is so low that a positive result is likely to be a false positive, and you would, in effect, tend to disbelieve the test result.

Again, for patient B, the reverse applies, and the PPV will be high.

Using imaging for screening

So far, we have been talking about using imaging (and other) tests to diagnose disease in patients with symptoms. This accounts for the major part of our practice in real life.

Screening, on the other hand, is a process where we look for disease in a group of people who have not yet developed symptoms.

Task 2: Go to box 2, and work through the questions.

box 2.

a) A number of requirements have to be met by a disease if a screening programme is to be feasible. List any that you can think of, before proceeding to the answers below.

b) List the desirable attributes for a screening test

Task 2 answers:

a) the conditions that must be met before mass screening for a particular disease can be viable are:

- the disease must pose a significant risk to health i.e. it would not be cost-effective to put in place a multi-million pound national screening programme for the common cold, even if it were possible
- the disease must have a pre-symptomatic phase during which it can be detected i.e. (to take a silly example again) there would be no point screening for broken legs, since there is no gradual transition from pre-symptomatic to symptomatic: you're OK one minute, and writhing in agony a second later, and so there is no window of opportunity in which to apply your screening test. Breast and cervical cancer are ideal candidates, because their natural history is well-understood; they are initially fairly slow-growing, and we have good tests (mammography and cervical smear tests) which can detect them.

- the disease must be common enough to make screening worthwhile i.e. if it is too rare, there will be implications for the accuracy of the test which would make screening non cost-effective (see task 3, below)
- there must be an effective treatment available there is usually little point spending a lot of money screening for disease if you can't do anything about it when you diagnose it (although there are exceptions in the case of some rare genetically-transmitted diseases, where it may be important for people in affected families to know that they have the condition or are carriers)

b) The desirable attributes for a screening test are:

- it must be widely available i.e. a test which is only available in a few centres would be no use for a mass screening programme
- it must be (relatively) cheap
- it must be safe no point using an imaging technique that delivers so much radiation that it causes as much cancer as it detects!
  (see http://www.cancerscreening.nhs.uk/breastscreen/risks.html)
- it must be accurate (and in particular, needs to be sensitive – see task 3)

Task 3: Go to box 3.

box 3

Taking the NHS Breast Screening Programme (NHSBSP) as an example, the expected number of cancers detected per 1000 women screened is about 4. Take this as the prevalence (0.4%) or prior probability (0.004) of the condition in this population.

Given this low prevalence:

a) what will the predictive value of a positive mammogram be? (high, medium, low)

b) and what will the predictive value of a negative result be?

c) you need a test with a high sensitivity in the screening situation, or you will miss many of the cancers you want to detect. What price do you pay for this high sensitivity?

d) the advantages of picking up disease in its pre-symptomatic stages are those of early diagnosis leading to more effective treatment, and better chance of cure. Can you think of any disadvantages to a screening programme like the NHSBSP?

Task 2

a) The

relatively high proportion of the positives will be false positives.

b) The NPV will be high – very few of the women actually have cancer, so most of the negatives will be true negatives.

c) The price you pay for high sensitivity is lower specificity: i.e. a higher percentage of false positives

d) Potential disadvantages of screening are:

- cost – there may be more useful things we could do with the large sums of money required for mass screening programmes
- dealing with the false positive results – for each true positive (early diagnosis, more effective treatment, happy patient) there will be several false positives. These women will be recalled for further imaging and possibly biopsy of abnormalities which turn out to be benign (unhappy, worried patient, more expense for the health service).

Screening and the public

The public tend to believe that screening must be a good thing under all circumstances, but they are often unaware of the potential drawbacks or of the need to confirm cost-effectiveness before launching a national programme. Even the established schemes have their critics: http://www.guardian.co.uk/uk_news/story/0,3604,245148,00.html.

There is also a lack of understanding of the limitations of any screening programme. In particular, there is a tendency (especially amongst politicians who should know better, not to mention the media) to assume that any false negative result must be the result of negligence. To take breast screening as an example, the signs of breast cancer on mammography are subtle and non-specific. For every suspect lesion seen on screening, a decision has to be taken, based on past experience, on the likelihood that it is malignant.

Set the index of suspicion too low, and although sensitivity will increase (fewer cancers missed), the number of false positives will rise to levels which cannot be coped with. Set it too high, and you will miss too many tumours. So, it's a compromise, and even the best of screening programmes will generate some false negatives, and women will present with breast cancer having been given the all clear on earlier screening.

This concept is not understood, and there have been examples in the UK where breast and cervical screening programmes have been vilified in the media for 'missing' tumours, when they have in fact been performing at a level which meets all the established criteria of acceptability.

A good example of this was the shock/horror press coverage when the cervical screening service in Leicester released the results of an audit of their programme in 2001. Not surprisingly, there were some cases of interval cancer (cancers arising in women after a normal cervical smear), and in some of these, re-examination of the smear showed abnormalities. The press stories were hysterical, and gave the impression that numerous women had died as a result of the incompetence of the screeners. In fact, the Leicester programme was performing well, and meeting all its targets. What journalists find impossible to understand is the concept that, as we have seen, no test is perfect, and that screening programmes have to make hard decisions about how many patients to recall.

For a more accurate, considered, news article, see:
http://news.bbc.co.uk/low/english/health/newsid_1310000/1310747.stm

ionising radiation: how dangerous is it?

The X-rays and gamma rays used in radiography, CT and nuclear medicine are examples of ionising radiation. As the name implies, when they interact with matter, they are capable of knocking bits off neutral atoms, leaving them with a net electrical charge. [refer to relevant bits of Tony's module]

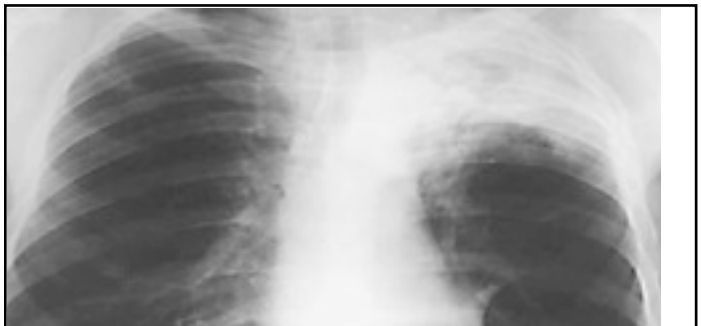
so what?

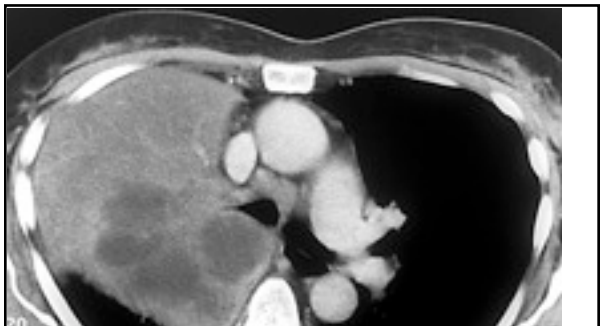
The interaction which matters biologically is that between the radiation and the nucleic acid in the nuclei of cells. The DNA (deoxyribonucleic acid) is the familiar double helix molecule of Crick & Watson which forms the genetic material of the cell. It is arranged into chromosomes (46 per cell) carrying the genes which code for the structural and biologically active proteins which make us what we are.

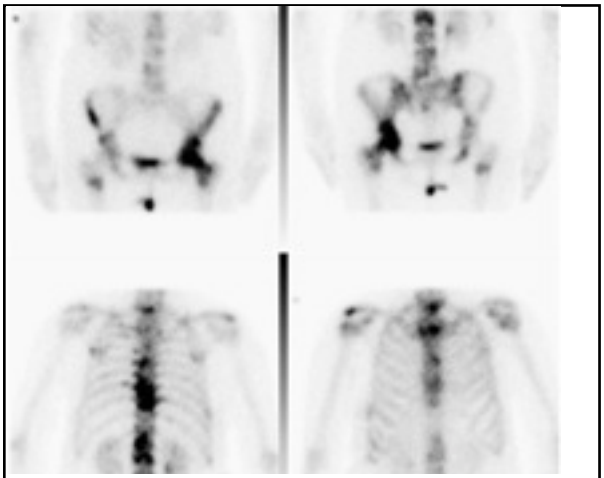
If cells are exposed to high enough doses of radiation, they will be killed. This is the basis of the use of radiation to treat cancer (radiotherapy). When smaller doses of radiation interact with DNA, the result can be subtle or not so subtle changes in one or more genes, producing mutations, but stopping short of killing the cell. How damaging this is will depend on which gene is altered, and whether it is altered enough to produce a protein which no longer does the job it was intended for. Ironically, although high dose radiation can be used to treat cancer, the most important effect of the low doses used in diagnostic imaging is the induction of tumours in the irradiated tissue.

radiation and cancer

The exact way in which radiation causes cancer is beyond the scope of this module (and the understanding of its leader). However, what seems to happen is that occasionally a mutation occurs in one of the many genes that have a regulatory effect on cell growth and division. The end result is a clone (line of cells arising from the division of a single precursor) of cells which divide in an uncontrolled fashion, and which lack the normal inhibitory mechanisms which prevent cells from infiltrating surrounding tissue. This tumour then grows and invades locally, and can spread to distant sites (metastasise) by way of the bloodstream or lymphatic system (see illustrations, below).







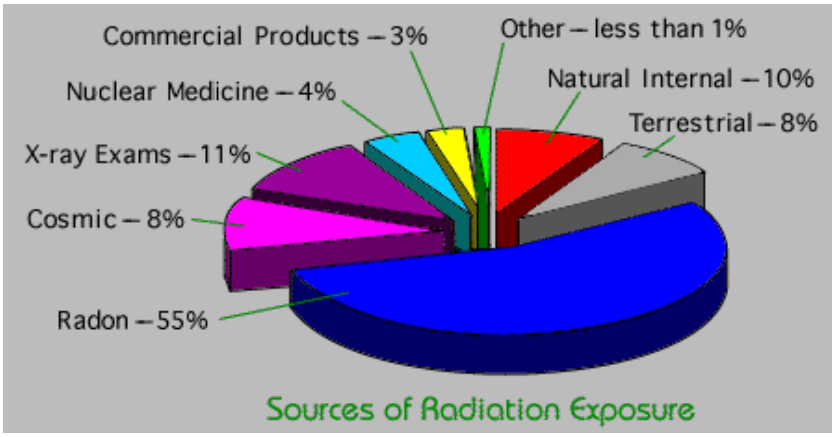
bone scan, showing widespread bone metastases from the lung primary

sources of exposure

It's easy to get caught up in discussions about the hazards of man-made radiation, medical sources in particular, and to forget that we are all being continuously exposed to radiation. The largest single component of this background radiation is radon gas arising from the rocks and soil beneath our feet, and often building up to quite high levels inside buildings. Other contributions come from cosmic rays and other natural emitters in the rocks and soil. The average annual dose from these natural sources is around 2.5mSv, about the same as the dose from radiography of the lumbar spine.

So why worry too much about the doses from diagnostic radiology? Well, although they only account for about 15% of the total dose received per year, medical uses constitute by far the biggest contributor to man-made (and therefore potentially avoidable) sources of radiation.

The pie chart below illustrates this. As an aside, note that the doses from all commercial sources, including nuclear reprocessing plants like Sellafield in the UK, are in the 'other' category, i.e. less than one percent of our annual exposure. This might come as a surprise, given the amount of media coverage generated by the emotive topic of nuclear waste discharges, and the numerous pressure groups clamouring for more expenditure to reduce the perceived hazard. A far more efficient dose-reduction measure would be to invest in providing our radiology departments with modern equipment. This is another example of the failure of the public (and the press and politicians) to take a logical view of risk assessment (see below, in 'communicating the risk).



just how dangerous <u>is</u> radiation in diagnostic doses?

The simple answer is that we don't know.

why don't we know?

because:

- the harmful effect we are looking for is very common in the population – approximately 1:3 of us will develop cancer at some stage in our lives
- although we don't know the exact size of the risk from diagnostic doses, we know it is very small, so it will never be possible to detect the additional few cancers caused by diagnostic irradiation against the high background level of 25-30%
- it is not possible to differentiate a cancer caused by radiation from one that would have occurred anyway
- the cancer will often not appear until many years after the exposure that did the damage

So, we assume the worst. Because, in theory, a single photon of ionising radiation could induce a critical mutation, the assumption is made that there is no threshold dose below which harmful effects cease to occur. And because we can't measure the effects of low doses directly, the known effects of high doses are extrapolated downwards, assuming a more or less straight-line relationship between size of dose and the magnitude of the effect.

There is a good summary of the current state of knowledge concerning the risks from low dose radiation at http://www.nrpb.co.uk/Press/Pr8-01.htm. Read it now.

communicating the risk to patients

How much information we should give to patients is a hot topic, and increasingly, the answer is 'all of it'. We are rapidly approaching the situation which obtains in the USA, where patients have to be informed of all the conceivable risks of any procedure or treatment, no matter how remote. Any attempt by doctors or other health care professionals to limit the information given is now seen as 'arrogant' or 'paternalistic', and is becoming as politically incorrect as the following cartoon (enlarge to view!):

SO: we are going to have to tell patients something about the risks of the radiation we use in diagnosing their disease. With the knowledge you have just acquired concerning the risks of low-dose radiation, have a go at the following exercises. These are designed to make you think about the issues involved in informing patients and obtaining consent – there are no right and wrong answers for many of the questions which follow.

task 4:

You have been charged with drawing up an information leaflet for patients attending your department for an X-ray examination. What do you think are the main factors which make this such a difficult task?

task 4 – answers:

These would be my answers:

- we are trying to tell patients how big the risk from low doses of radiation is, when we don't actually know ourselves! It is very difficult to explain the uncertainty to scientifically unsophisticated patients without making them think we are either incompetent or hiding something.
- the general public (including you and me!) have a very idiosyncratic approach to risk. We don't simply look objectively at the magnitude of the risk in question, and then make a judgement. Risk perception is

coloured by all sorts of cultural and psychological factors, and has become a fascinating area of research. There is a list of some of these factors in the box:

task 5:

Using the information in the box, explain why it is that a patient who is a heavy smoker, and has developed lung cancer as a result, can still claim to be worried about the adverse effects of the radiation used for the CT scan used in staging his tumour.

some of the factors modulating risk perception, based on Covello and Merkhofer (1994):

- catastrophic potential – people are more concerned about fatalities and injuries that are grouped in time and space (aeroplane crashes) than about fatalities and injuries that are scattered or random in time and space (car accidents);
- familiarity – people are more concerned about unfamiliar risks (ozone depletion) than familiar

task 5 – answer:

the smoking risk is one which he took on voluntarily, and is therefore under his own control. He also feels that he understands the risk from cigarettes, which is a familiar one.

Radiation is an unknown quantity, and is something which invokes dread in the public. Also, it is being forced on him, and therefore out of his control. We are also asking him to trust someone else (i.e. us) not to give him a dangerous dose of radiation.

but it's not all bad news!

In all of this, we have been concentrating on the risks of radiation. There has so far been no consideration of the benefits of making a diagnosis and giving the patient the appropriate treatment. When explaining risks to patients, they must therefore be put in the appropriate context. We should only subject patients to ionising radiation if the likely benefits to their health outweigh any small risk. This is the basis of the justification of medical exposures; a concept which is at the heart of the legislation governing our use of radiation in medicine (q.v.).

task 6:

In note form, sketch out how you would draw up an information leaflet to put the radiation risks into perspective for patients. How would you set the leaflet out, what section headings would you use etc.? In each section, indicate the approach you would use. No one is marking this, it's for your own benefit, but do go through the exercise as if you were doing it for real.

task 6 – answer:

There is no answer – it's all very difficult! However, follow the link to a recent attempt to produce such a document (authored by the NRPB, CoR, RCR and RCGP): http://www.nrpb.co.uk/Miscpubs/X-ray.pdf

What do you think of it? Could you have done better?

Legislation and ionising radiation

Given that ionising radiation is potentially hazardous, it's not surprising that Health and Safety legislation has something to say about the way we use it. Although there are a number of separate pieces of legislation covering various aspects of practice, the two main set of regulations in the UK are:

the Ionising Radiation Regulations 1999 (IRR)
and
the Ionising Radiation (Medical Exposures) Regulations 2000 (IRMER)

The first of these covers the exposure of radiation workers and members of the public, the second deals with exposures to patients. As this module is about clinical aspects of imaging, I'll be concentrating on IRMER in what follows.

history: Until 1988, there were no formal regulations concerning the exposure of patients. For many years, IRR had set out dose limits for the staff of radiology departments (and other radiation workers) and for members of the public. There were no such limits for patients, on the grounds that these might compromise our ability to reach a diagnosis. Patient exposures were, however, subject to the ALARA principle (As Low As Reasonably Achievable), more recently known as ALARP (As Low As Reasonably Practicable). In other words, doses to patients would be kept as low as possible consistent with reaching a diagnosis, always assuming that any radiation risk was outweighed by the potential benefit to the patient's health.

In 1988, the Protection of Persons Undergoing Medical Examination and Treatment regulations, ever after known as POPUMET, came into existence. For the first time, these laid down guidance on the exposure of patients. We needn't go into any detail about POPUMET, except to say that, although the legislation was a good first attempt at regulating patient exposure, there were a number of failings which were addressed to a greater or lesser extent in IRMER, which replaced POPUMET in 2000.

IR(ME)R

These regulations take the logical approach of defining the key roles undertaken by medical and paramedical staff when patients are exposed to

radiation for diagnostic purposes (the Regulations also cover the therapeutic uses of radiation, but that's outside the scope of this module).

Task 7

The legislators identified 3 key roles in the exposure process. Look at the pdf copy of the legislation at the following web address, and then list those key roles with the responsibilities that each of them carry.
www.doh.gov.uk/irmer.htm

Task 7 – answers

Before a procedure is performed, someone has to request it, someone else should, ideally, confirm that the request is reasonable, and then a third person or persons has to carry it out. IRMER identified the following roles:

the referrer This is the person, usually medically qualified, who decides that the investigation is required, and sends the referral card or letter to the radiology department.
the practitioner Not the word that most of us would have chosen, because we already have lots of 'practitioners' in our hospitals (e.g. medical practitioners, radiographer practitioners, nurse practitioners etc), but this is the key role in the whole process. The IR(ME)R Practitioner's only role is that of justification of the exposure. In other words, they have to examine the clinical details on the referral card and decide whether there is sufficient reason for exposing the patient to radiation. If they are happy, they authorise the exposure, if not, they refer it back to the originator explaining why the test has not been authorised, and possibly requesting more clinical information.
the operator The obvious example of an operator is the radiographer who presses the button to make an exposure, or the member of staff who injects a patient with radioactive material for a nuclear medicine scan. However, the definition is much wider than that, including anyone who performs any practical task related to the exposure. For example, the physicist who calibrates the equipment used in making the exposure is an operator.

task 8

Go back to the IR(ME)R website, and answer the following questions:

a) who is responsible for ensuring that a department complies with the requirements of the legislation?
b) what are the responsibilities of the referrer?

Task 8 - answers

a) the employer is responsible. This is important, as it ensures that the legislation is taken seriously by Trusts.

b) the referrer is responsible for supplying the Practitioner with sufficient clinical information to justify the procedure requested.

Quality: of the image and of the service

So far, we've been talking about the accuracy of clinical testing in a fairly general sort of way, but this course is all about imaging, and so we have to at least mention the concept of image quality. Doesn't matter how powerful a particular imaging test is in diagnosing a particular condition, if the images we produce are rubbish, the test, in our hands, will be useless.

Not surprisingly, radiologists, radiographers and medical physicists spend a lot of time measuring and discussing image quality. Many of you will be dealing with these issues in some depth in other modules, but this module is supposed to be about medical imaging in clinical use, and so what I want to do here is to look at the topic of image quality from a clinical viewpoint.

To demonstrate what I mean, go to the following website. I don't want you to read all the information there in any detail, but just look at the boxes at the top of the document which list the contents:

http://www.bh.rmit.edu.au/mrs/DigitalRadiography/DRNotes/iq/

defining image quality

Ask a radiologist for a definition, and he or she will probably start by listing the sort of parameter you found in the contents of the web page you just visited. These are all fairly 'hard' physical measures of image quality, and it's interesting that almost everyone you ask about image quality starts off by actually talking about the methods we use to measure it. That's quite a different thing to defining it, of course, and after a bit more thought, our radiologist will probably say that it all depends on what you want to achieve with the help of the image in question. In other words, there is no absolute measure of image quality, at least from the point of view of the end user.

For example, a 'poor quality' portable chest radiograph on a sick elderly patient may be good enough to confirm or refute a clinical suspicion of cardiac failure, but it probably wouldn't be adequate to assess the presence, absence or extent of interstitial fibrosis in a patient with suspected diffuse lung disease. This has been summed up in the statement that the aim of diagnostic radiology is to 'obtain images which are adequate for the clinical purpose with the minimum radiation dose to the patient', and it is that word 'adequate' which is the key to any discussion of radiological image quality.

when it comes to quality, do we always need the best? There has been a tendency to assume that we should always strive for the highest possible image quality (however defined), on the basis that to accept anything less would represent a dereliction of our duty. However, given the inevitable mismatch between our aspirations and the resources available, we need to ask whether it makes sense to use our money in an attempt to squeeze out as many line pairs per millimetre as possible from the available technology, when it may not be necessary for the clinical task at hand. One example is the assertion that 2000^2 (or even 4000^2) monitors are required for satisfactory image display in the digital environment. There is good anecdotal and experimental evidence that this is not the case [http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?CMD=Display&DB=PubMed]. This paper showed that, while radiologists 'liked' the higher resolution display, it actually made no difference to what they could see, or to the accuracy of their reports. This matters, because the additional expense involved in buying such high resolution monitors could well result in the cost of a digital network becoming prohibitive.

There are therefore good clinical and operational/economic reasons for thinking seriously about image quality, and what we mean by the term, and so we do need to be able to measure it, even if we have trouble defining it in absolute terms. There are different sorts of measurement available, and it may be that looking at these methods will give us a clearer idea of what we mean by image quality:

physical: this is the sort of thing that we found on the website. Examples include contrast, latitude, MTF, detective quantum efficiency (DQE), line pairs per millimetre, and there are many others. These 'hard' objective measures are reproducible, and allow direct comparison of the performance of different imaging systems or techniques, but they take no account of the role of the observer in the diagnostic process. There's no time here to go into the psychology of perception in any depth, but for amusement, and a very brief introduction to the subject, go to:

http://www.expasy.ch/UIN/html1/projects/MedArtImages/MedArtImage.html

(and check out the pictures!)

and to get a hint of how complex the underlying science of how we see things can be, roam through the following site (and at least have a look at the optical illusions!):

http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip--3.html

subjective: Attempts to perform realistic assessments of image quality in clinical practice frequently involve the use of subjective rating scales. A literature search reveals that most studies reported in the clinical literature use this approach, either alone or in tandem with more objective measures. Below is an example of an image rating sheet used in Leeds to evaluate the performance of a new digital chest radiography system in clinical use.

area	1	2	3
lung fields	☒		
bone detail - ribs		☒	
sub-diaphragmatic			☒
costo-phrenic angles		☒	
trachea & main bronchi		☒	

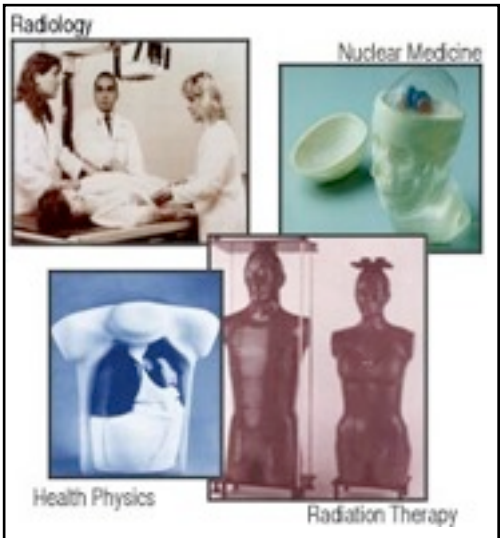
where: 1 = good
2 = adequate
3 = poor

Although this sort of 'soft' assessment relates to real life and allows us to rate a particular imaging system or technique on the strength of its performance in a clinical environment, the reproducibility of the results and their relevance to other users of the same technique will always be open to question. It is also true that the radiologist's subjective perception of the relative quality of two different sets of images may not relate to the level of his or her performance when using those images in a clinical setting (as we saw earlier with the paper comparing medium and high resolution monitors).

objective/realistic: It is possible, but not easy, to develop methods of assessment which retain at least some of the realism of the subjective methods, and therefore have credibility with end users, while at the same time employing scientifically respectable methodology, and producing data which are reliable and reproducible. Studies using very simple equipment, such as the Leeds threshold contrast detail detectability (TCDD) test object, while not producing anything resembling a clinical image, do measure a clinically important image parameter (the ability to pick out structures

against backgrounds of slightly different density), and allow a group of observers to compare the performance of different imaging systems in a way which allows for human perceptual variability, and produces results which can be used to make rational decisions about the application of new technology.

Studies using more complex phantoms (see box) can closely reproduce the conditions applying to some clinical tasks; for example, a good chest phantom with added 'lesions' of varying size and contrast can provide a rigorous and realistic test of a chest imaging system, and the same is true in other areas, such as mammography. In other cases, it is possible to use a well-defined clinical condition, preferably one which lends itself to grading based on the imaging findings, for example, the plain film changes in the hands of patients with hyperparathyroidism: http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?CMD=Display&DB=PubMed



Phantoms are models made of tissue-equivalent material which contain 'lesions' of various sizes and densities. These can be used to mimic patients in experiments to assess the image quality produced by different radiographic techniques.

image quality and outcome

There is currently a lot of interest in medicine generally in devising **outcome measures** for the things we do to patients. That is, we can only assess how well we are performing if we first define the desired outcome of whatever it is we are doing, and then devise some way of measuring how successful we are in achieving that aim. But it's not as easy as it sounds. In radiology, one outcome measure for, say, chest radiography, would be that we produce a film of diagnostic quality and interpret it correctly. That's a pretty limited outcome measure, but even that can be hard to obtain.

task 9: What problems would you anticipate in auditing performance using the outcome measure for chest radiography outlined in the previous paragraph?

task 9 - answer: possible issues are;

- how do you define 'diagnostic quality'? Not insoluble, and it would require the reporting radiologist to assess this at the time of reporting, bearing in mind that this will be a relative measure, depending on the reason for performing the x-ray. An alternative would be to have the films for the audit assessed by an independent radiologist, but all this takes time and people – both commodities being in short supply
- how do we decide whether the film was reported correctly? There is unlikely to be a 'gold standard' against which you can assess the accuracy of the report. It's not as if the patients will all be going straight to the operating theatre to have the diagnosis confirmed. So, you will probably end up by following the patients up over a period of time, and seeing if their progress, or lack of it, suggests that the radiologist was right. Again, lots of time-consuming work for someone.

Ideally, we need measures which actually look at the outcome for the patients: in other words, what effect does our intervention (in this case, the chest x-ray) have on the management and outcome of the patient's disease? In many areas of practice it is difficult or impossible to devise sensible ways to measure this.

It is possible in some instances though, particularly in areas like interventional radiology, where a single intervention is undertaken with a specific aim. For example, where the radiologist is undertaking angioplasty (see above) of the femoral artery in a patient where narrowing of the vessel

produces pain, and limits their activity, there are several outcome measures which can be employed. The obvious one is the impact on the distance the patient can walk. If they were brought to a halt by pain after 50 yards of slow walking before the treatment, and post-procedure they can go hiking in the Yorkshire Dales, that's an excellent result. Other outcome measures can look at the morbidity and mortality of the investigation. For example, you can record the incidence of bleeding from the arterial puncture site in patients undergoing angioplasty. However, in most cases we are some way from devising any sensible outcome measures with which to judge our performance.

The quality of the service

Although image quality is clearly important, there are a lot of other factors which contribute to the overall quality of the imaging service.

task 10: List the factors, apart from image quality, which you think are important in the provision of a quality imaging service.

task 10 – answer: No right and wrong answers to this, but here's my list:

- Access: increasingly, this means providing a comprehensive service during office hours, with arrangements for adequate out of hours cover. There is increasing pressure for a 24 hour a day, 7 day a week service (see unit 1).
- Timeliness: no good having state of the art equipment and good staff if patients have to wait so long that they are either dead or recovered by the time their appointment comes round. Equally important is the prompt delivery of the report to the referring clinician.
- Physical environment: the department should provide pleasant changing and waiting accomodation for patients and their attendants. This has historically been a problem due to the ageing building stock in the NHS. Slowly being addressed as new capital schemes get off the ground.
- Monitoring the quality of care, with commitment to audit and a process of continuous improvement
- Protection of the patient – full compliance with Ionising Radiation legislation.