The Logical Analysis of Clinical Medicine, Part One

The Scientific Basis of Medicine

Presented by Professor Wilfred Card, University of Glasgow.

University of London Audio-Visual Centre, 1975.

Made for British Postgraduate Medical Federation.

Illustrations in the programme are reproduced by permission of the following:

The Wellcome Trustees

The Royal College of Physicians

William Collins, Publisher

Glasgow University TV Service.

Graphic design by Janice Minton and Hank Young.

Produced by David Sharp.

Black-and-white

Duration: 00:38:23:06

00:00:00:00

<**Opening titles**>

< **Unnamed narrator quotes from Roger Bacon over image of Roger Bacon**>

"Mathematics is the key to the sciences."

< **Unnamed narrator quotes from Gottfried Wilhelm Leibniz over image of Gottfried Wilhelm Leibniz**>

[n.b. Leibniz is spelled incorrectly as Leibnitz in the image]

"There would be no more need of dispute between two philosophers than between two accountants. It would suffice for them to take their pencils in their hands, sit down at their slates and say to each other 'Let us calculate.'"

<**Unnamed narrator quotes from Karl Pearson over image of Karl Pearson**>

"I believe the day must come when the biologist will, without being a mathematician, not hesitate to use mathematical analysis when he requires it."

<**Card to camera**>

If we are to develop a mathematical structure of medicine, we must define our objective and no text book seems to discuss this. Hippocrates said:

<**Unnamed narrator quotes from Hippocrates over image of Hippocrates**>

"I would define medicine as the complete removal of the sick, the alleviation of the more violent diseases and the refusal to undertake to cure cases in which the disease has already won the mastery, knowing that everything is not possible in medicine."

<**Card narrates over image of Hippocrates**>

These sentiments are admirable, but this objective is impractical since it does not take into account cost. In Britain, we have 55 million people and an NHS budget of about 3 billion.

<**Card to camera**>

Our resources are limited. How can we deploy them to the best advantage? If we are to do this, we shall need to use decision theory. Here's an outline of a very simple decision system.

<**Card narrates over animated diagrammatic chart**>

Here is the patient and there is a choice of two treatments. To each treatment there is a choice of outcomes. We shall have to attach a measure of worth or value to the outcome and we use the word ‘utility’ for this measurement. Since the doctor has no certainty but only a probability of achieving this, it is more accurate to speak of expected utility and hence of a doctor’s objective as the maximisation of expected utility.

<**Card narrates over image of Von Neumann**>

The person responsible for this decision theory was the great Von Neumann.

<**Card narrates over animated diagrammatic chart**>

In this logical analysis of clinical medicine, all that can be put forward is a very tentative scheme. We shall need to think of our patient as a set of symptoms, signs, radiological and laboratory data, each of which is a defined element in our system. It’s useful to have a name to describe any of these.

<**Card narrates over page of old medical text book**>

I.J. Good suggested the word ‘indicant’, an old medical word with its sense slightly altered.

<**Card narrates over description of ‘indicant’**>

We’ve defined it here as any piece of evidence that might be relevant to the probability that a disease is present.

Though the definitions of all the indicants can be very difficult, in our experience it has never proved impossible.

<**Table**>

Nocturnal pain

Patient wakes up with pain and takes milk, a snack, antacids, or a little warm water and gets relief.

Never = 1

Rarely = 2

Often = 3

<**Card narrates over above table**>

Nocturnal pain we've defined in this way in a study of dyspepsia.

00:04:30:00

<**Card to camera**>

The procedure we use to elicit an indicant we call a test. So that the question is a test, examination for clubbing is a test, an x-ray is a test and so on. Of all the methods of examination, history taking will be needed for a very long time because it provides very early and sometimes very strong evidence of disease so that we can often, as you know, make the diagnosis on the history alone.

<**Card narrates over sliding diagram**>

History taking starts with an event in the body. Here is the gut contracting. Nervous impulses pass to the brain / mind of the patient where an experience takes place and is coded into a signal which is received and decoded by the doctor, interpreted and he records 'colic.'

<**Card to camera**>

In this chain, though, there are many weak links and there's only a degree of probability that an event is correctly recorded. The analysis of such probabilities leads to the estimation of errors, as we shall see later. Some of these difficulties are emphasised when we use interrogation by computer.

<**Unnamed narrator narrates over a moving image of a man in front of a computer screen**>

Here we see a patient being interrogated by a computer. The questions are typed onto the screen and the patient answers by pressing one of the three buttons which are marked ‘yes’, ‘no’ and ‘don’t understand’. The computer then types out the next question, the choice of which depends on the patient’s answer. One of the first things we realised about this form of interrogation was the importance of ensuring that the patient understands the questions. We therefore test each medical term used to see if, without context, it is understood by 95% of patients. In one survey we found, for example, that only 71% of patients understood the term ‘heartburn’ and so we now have to add “I expect you know what heartburn is. It is perhaps best described as a burning feeling below the breast bone. Tell me, do you suffer from heartburn a lot?” We also try to use very simple phrasing but if there’s still apparent uncertainty the patient can press the ‘don’t understand’ button and an alternative wording of the question will be presented.

Many patients find the strict discipline imposed by ‘yes’ and ‘no’ answers quite acceptable and satisfactory. But others have commented that they would like to be able to qualify their answers. We therefore designed this keyboard which can be switched for suitable patients to a 7-button keyboard to allow the qualifications ‘certainly’, ‘probably’ and ‘possibly’ for ‘yes’ and ‘no’ answers.

<**Unnamed narrator to camera**>

We have tried to evaluate computer interrogation with respect to intelligibility to the patient, accuracy and cost. The objective attitude-scaling techniques we have used have shown that 82% of patients have favourable attitudes towards computer interrogation. That’s about the same proportion as the favourable attitudes towards interrogation by a doctor. Up to 49% actually have more favourable attitudes towards computer interrogation than they have towards medical interrogation with a doctor.

Now, since interrogation by a computer is also comparable to a consultant with respect to accuracy and cost, we feel this method must be the future.

<**Card to camera**>

Given the sufficient data there is, there is nothing to stop a computer programme from calculating the probabilities of the diseases indicated while the patient is being interrogated. While the simplest responses of the patient is through keyboard, there might be times when it is useful for the computer to recognise sufficient speech to respond to 'yes' and 'no.'

<**Card narrates over moving image of a man talking on a telephone**>

And speech-pattern recognition has now been developed to such an extent that this is possible for these simple replies. It therefore becomes possible to interrogate and receive replies through a telephone conversation.

<**Moving image of man giving simple answers into a telephone**>

00:09:43:00

<**Card to camera**>

This process of definition of symptom indicants has to be extended to physical signs, radiological signs and so on.

< **Table**>

Pointing sign

Positive if one or more fingers used to indicate the area of pain

If one = 2

If more than one = 1

Negative if flat of hand used or hand movement occurs around an area = 0

<**Card narrates over above table**>

Here is an example of a pointing sign used in dyspepsia and this is how we've defined it. And it would obviously be very useful to have nationally or even internationally agreed definitions as we have broadly speaking in dyspnoea.

<**Moving table showing 5 grades of dyspnoea**>

1. **No abnormal breathlessness**
2. **Able to walk normally without breathlessness on the level, but breathless on hurrying or climbing slight hills.**
3. **Able to keep on walking at own slower than average pace on the level**
4. **Forced to stop for breath when walking at own slow pace on the level**
5. **Breathless on slightest exertion such as washing or undressing.**

<**Card narrates over above table showing 5 grades of dyspnoea**>

There are five grades of dyspnoea which are widely recognised and this is the kind of definition that we really need throughout medicine.

<**Card to camera**>

Having specified the patient in this way, the next step is to allocate him to a disease class. What do we mean by a disease? Most doctors think of a disease as an abstraction, that is as a mental concept, drawn from a number of examples from a number of patients. This is what is known as the nominalist view, philosophical view, which is two thousand years old and it might be summed up, perhaps, in the aphorism of our medical teachers: "diseases do not exist, only sick people."

So that one question we might ask, and the mathematicians ask this, are disease classes really necessary?

<**Card narrates over animated diagrammatic chart**

Could we not assign the patient to a treatment class as being shown here without going through the intermediate step of assigning him to a disease class? Now, in general this doesn't seem possible because of the size of the state of health vector – the xn can be very big and the number of treatments. We should have to have a function, treatment as a function of the patient's state and we seem to need these disease classes for economy of thought. Some of these disease classes contain logically defining characteristics.

<**Table**>

> **Typhoid fever is present if blood culture is positive during the first week of the disease, the blood culture being repeated a second time if the first is negative (p=0.99)**

<**Card narrates over above table**>

One of these disease classes is possibly typhoid fever which can be defined in this way. But clearly we encounter a number of conditions which most doctors would, diseases which don't have defining characteristics and cannot be defined in this way and can only be defined statistically.

<**Table**>

> **Depressive syndrome**
> **Weight loss constipation, early wakening, depressive psychomotor activity**

<**Card narrates over above table**>

An answer might be the state of depression and here we have a number of symptoms which may or may not be present but we seem to say a state of depression exists or is present provided by several characteristics.

<**Card to camera**>

So the problem becomes that of defining a disease class on some statistical basis and this is a general problem of classification which exists throughout all of science.

<**Card narrates over diagrammatic chart**>

In disease, the space in which disease exists is the event space, the space on the left, in the body. But of course we have no direct access to that except through some information channel, and the only space we can deal with is the indicant space, the space of appearances, which is shown on the right. The black lines indicate where an event has come through correctly. The white lines show where error has occurred. E1 can come through as I bar 1 (bar is negative), E3 comes as I bar 3 and so on, E5 can come through in two forms, E6 never comes through at all.

00:14:06:20

<**Card narrates over a further diagrammatic chart**>

Suppose in the very simplest case, the patient could be characterised in two dimensions. If the characters were continuous variables, you can imagine that a given patient could be represented by a point in this two-dimensional space, and the two groups corresponding to the two diseases might separate out. But of course in our disease space, we're dealing with a multi-dimensional space and we can't show you that but these photographs may perhaps give you some idea of the kind of problem.

<**Card narrates over abstract still images**>

The problem here is – is this four clusters, four diseases or is this one?

<**Card to camera**>

So the problem is one of analysing these clusters – how can we do this? Suppose we mention that the character is either present or absent

<**Card narrates over chart showing movable plus and minus signs**>

And that here we show ten patients, each with five indicants, which are either present or absent, and they have this set of characters. If we sort them, and this is of course what the computer would do, we shall find that any pair in the top group has three characters in common and any pair in the bottom group has three characters in common. But no example in the bottom group has more than two characters in common with the top group, so that we've separated out these ten patients into two clusters of five. And this is the general kind of principle of cluster analysis.

<**Card to camera**>

Most workers in the field, in the logical analysis of medicine field, have very reasonably chosen disease classes which can be clearly defined and the techniques that I have outlined have only been used rarely. We try, as an example, to see if proctocolitis could be separated from Crohn's disease of the large bowel.

<**Card narrates over diagrammatic chart**>

And the separation, which is of course in a multi-dimensional space; we used 107 indicants, 109 patients, and we got a number of clusters, the big cluster on the left was almost entirely composed of proctocolitis. And the Crohn's disease patients appeared as small clusters as kinds of satellites. The general conclusion we can come to is that it looks as if we can define the disease class in this formal way.

<**Card narrates over a further diagrammatic chart**>

The outline of the general objective of general management is to maximise expected utility, but it's usually necessary to break down the process into the initial step of assigning the patient to a disease class. They call this diagnosis in the restricted sense. If all these diseases contained logically defining characteristics, we could do this biological tree.

<**Card narrates over an open book displaying a biological tree**>

This is the method of identifying an unknown bird, or in this case, an insect. Can the insect fly? Yes/no, and you make the appropriate step, which is a move through a decision tree. It's a very old method and the botanists have used it for several hundred years.

<**Card to camera**>

Construction of such trees is not simple, but it's now possible to construct an optimal tree, optimal in the sense that it allows identification on the average in the fewest possible steps. But diagnosis in this way is not usually possible and has to be probabilistic. The process we call mathematical inference in the way in which degrees of belief are altered by data and there are various ways of doing it.

<**Card narrates over animated chart showing mathematical equations**>

The method which has been most widely used is called Bayesian Inference and what the doctor wants to know is, given the evidence, what is the probability of a given disease? He has, to calculate this, two kinds of information – the initial probability of disease, how likely it is to turn up and the probability of the evidence, given the disease. These two can be combined in this simple way, the product of the disease, given that the evidence varies, is the product of the other two probabilities and can be turned into a true probability by using rather complicated, normalising factor in the denominator.

00:19:58:00

<**Card to camera**>

This is only one step, and as each piece is collected, the probability of a given disease can be re-calculated. We can see this in an example of a patient with jaundice.

<**Table**>

J.C. a 45 year old man
six weeks diarrhoea and painless jaundice
later severe itching
no anorexia or weight loss
no ascites or liver enlargement
taking librium

<**Card narrates over above table**>

This is the sum of the symptoms and signs of a man with jaundice.

<**Table**>

Initial	**Bilirubin**	**4.0mg**
	Alk Phos	**32 KAU**
	SGOT	**39IU**
	WBC	**8400**

<**Card narrates over above table**>

Some of the laboratory findings.

<**Card narrates over animated graph**>

And when these are introduced one by one and the probabilities calculated, you will see the way in which the probability of cancer arises until we get to test 5 and then the history of taking drugs is brought in and this completely alters the probability and the probability of drug-induced jaundice becomes very high.

<**Table**>

Case J.C.

predicted by probability odds

drugs	**.985**	**66**	**-1 on**
cancer	**.004**	**250**	**-1 ag**
necrosis	**.002**	**500**	**-1 ag**
stones	**.002**	**500**	**-1 ag**
hepatitis	**.001**	**1000**	**-1 ag**
etc			

<**Card narrates over above table**>

This table shows the probabilities and the odds – and you'll see the probability of drug jaundice was 66-1 and that was, in fact, the correct diagnosis.

<**Card to camera**>

You'll see the process is sequential, a series of steps and the experienced clinician selects each step. Given a set of diseases at the back of his mind, he chooses the test which is most likely to be profitable. It is possible to do this formally, to calculate the choice of the next best test.

<**Card narrates over animated graph**>

This is possible and we can get a measure of uncertainty. And in this illustration, you see the uncertainty as a graph which diminishes during the diagnosis of jaundice in the last illustration. The dots are calculated, expected uncertainty from the test, and the continuous lines are the actual uncertainty. So it is possible, at each step, to calculate the test of greatest expected informativeness. What we may colloquially call ‘the best bet.’

<**Card to camera**>

If we’re going to do a block of tests, for example biochemical tests, this doesn’t matter, but if we have some expensive or elaborate tests and we’re going to include costs, then it may be very useful. While there are other methods of inference, the basic method has proved powerful and robust, is computationally straight-forward and lends itself to the sequential choice of tests. When we say the method has proved robust, we mean it seems able to overcome its apparent weaknesses. The most obvious weakness is that though the data is appendant, the assumption is usually made that they are independent; we enlarge on this a little later.

Now, all doctors, however experienced, are liable to error and if we want to formalise medicine we need to analyse this business of making mistakes. We want to know how to estimate our error rate, we want to know how important the error rate is in affecting the evidence we elicit. First, how can we analyse error?

<**Card narrates over animated chart**>

One way is to look at it like this. Supposing that an event, E, either happens or doesn’t happen, E bar. Given that the event is a real thing – the patient really has heartburn or an enlarge spleen and so on, and if it happens and you assert that it doesn’t happen, F bar, this is the error of the first kind or false negative error, alpha. There’s a certain chance or probability of your asserting this. Since you must either assert that this happens or doesn’t happen, the probability that you are is 1-alpha. Similarly suppose that the event doesn’t occur, E bar. If it does, then your error is the false positive error, or error of the second kind, beta. Since as before you must either say that it occurs or doesn't occur, you'll be correct with probability 1-beta.

<**Card narrates over mathematical table**>

So that if we think of this business of making mistakes as an error rate in an information channel, we can put these rates into a table, the mathematicians call it a matrix and this matrix completely describes the channel. You can guess how this process could be extended to conditions where there are three or more probabilities.

<**Card narrates over further mathematical table**>

This is an example of the perfect doctor, when E1 occurs, he writes down F1, when E2 occurs, he writes down F2 and so on, he never makes a mistake, and this, in the matrix algebra, is a unitary matrix.

00:25:37:00

<**Card to camera**>

Sometimes a distinction is drawn between what is called soft data and hard data; soft data being obtained by history-taking and hard data is the data supplied by laboratories. But the important distinction is surely that of error rates. This is a useful way of thinking about error and how we are going to estimate these error rates. There's an obvious difficulty here when dealing with symptoms because for quite a lot of the time we don't know and have no means of telling whether the patient has the symptom or not. Now, if several doctors independently question the patient about that symptom they will agree or disagree and it is possible to devise a model which can estimate their individual error rates from the extent of their agreements and disagreements.

<**Card narrates over mathematical tables**>

This shows the agreement – 1 is yes, 0 is no, four possible kinds of agreement and disagreement; and with three doctors there will be of course eight possible combinations. We've tried this in a model making the assumption that alpha = beta and the error rates of three consultants are shown and the fourth consultant was the interviewing computer. The computer had a significantly greater error rate than did the doctors.

In some circumstances we shan't be too badly out if we assume the two kinds of error rate equal as here. But there are situations in which the false negative and the false positive rates are inversely related. For instance in the diagnosis of malignancies in cell smears.

<**Card narrates over graph**>

And this is shown on a slide here and it produces this kind of hyperbolic relationship between the two kinds of error.

<**Card to camera**>

Error rates can be very low indeed, for instance the assertion that there is a fracture in a long bone, and error rates can be very high when found with certain symptoms in dyspepsia. What we want to know next is what are the effects of error rates on the evidence we are trying to collect. Here we can only discuss the very simplest situation.

<**Card narrates over mathematical tables**>

First we have to have some measure of the weight of evidence and we can measure this by the change in the probability that the weight of the evidence produces but it turns out it's simpler to turn the probability into odds, as shown on the illustration. And the ratio for the change of odds, after you've collected the evidence, gives us a measure of weight of evidence.

As an example, in our gastrointestinal department, the odds against carcinoma of the stomach of a patient walking in, would be at least 50-1 against. Some time ago, a patient came in, sat down and eructated very foul gas. This must mean stenosis, probably with achlorydia, almost certainly with carcinoma of the stomach and the odds must go up, somewhere near 100-1 on, the odds have gone up 5000 times. The logarithm of 5000 is 3.7, the units are bans and this piece of evidence which is very great would be 370 centibans, if those particular odds are correct.

We can now extend the concept a little further to expected weight of evidence, that is the weight of evidence we can expect to get on the average from a particular test in a particular disease. And we can go further and plot the loss of the weight of evidence against the error rates. We've made the assumption here that alpha = beta and alpha + beta is plotted at the bottom and the percentage loss of expected weight of evidence is the ordinate. When alpha + beta = 1, no information, no evidence can get through and there's 100% loss.

<**Card narrates over humorous picture of a missionary and a native**>

This reminds one of the island in the Pacific where there were two kinds of natives – the goodies and the baddies and the goodies always told the truth and the baddies always told a lie. The first native the missionary met he asked if he was a goody and he said 'yes' and he was no wiser. Alpha + beta = 1.

<**Card returns to previous mathematical table**>

The part of the curve that we're interested in is the left-hand part of the curve and you see that where the error rates are both .2, and this would occur quite frequently in clinical practice, we lose on the average more than 60% of the weight of expected evidence, it also follows that if we could get rid of these high error rates we should get a lot more evidence. And you can think of examples in medicine where this has occurred.

00:31:28:22

<**Card to camera**>>

Now whatever the mathematics of the neuronal network up here, it is undoubtedly a superb mechanism for recognising patterns. And patterns can be of different kinds. There are auditory patterns and musicians are extraordinarily skilled at recognising a piece of music and there are visual patterns.

<**Card narrates over moving image of boy sorting letter cards**>

Here is a child trying to sort A's, B's and C's in different type fonts, upper and lower case. It's able to sort them correctly and that's something I think no computer can do yet.

<**Card to camera**>

The patterns can be medical. Try this one.

<**Card narrates over animated table**>

Wasting, jaundice, scratch marks, area of pigmentation in upper abdomen … the diagnosis is … ?

<**Card to camera**>

Carcinoma of the head of the pancreas. A lot of you will have recognised that pattern, literally, in less than a second. I think we should all agree that a lot of medicine falls into this pattern-recognition business. Once you've seen a striking pattern, you can remember it, even 20 years later. How can we analyse this business? I'm going to begin to do this in the most elementary way; we're going to see how far we might explain some of it, using the concept of dependent probabilities.

First, what do the statisticians mean by independence?

<**Card narrates over moving image of deck of playing cards and dice**>

If I throw a dice, since there are six faces, the chance that I throw a 1 is one sixth. If I draw a card at random, the chance that I draw an ace is one thirteenth. What is the probability of the joint event, 1 and an ace? It's the product of the two and if this holds because we believe that one does not affect the other, then we say the two events are independent.

<**Card to camera**>

If they are not independent they are said to be dependent. What's the chance that any out-patient has a raised pulse-rate? Not very great. But suppose we know already that the patient has a raised temperature, then the chances will be very great as these two events are strongly dependent. Now where the dependence occurs throughout the whole set of diseases it may not be very interesting except that a strongly dependent event may not contribute any new evidence. But where the dependent is within one disease and is not within another, then we may get the basis of a pattern that may be recognised. Look at this model.

<**Card narrates over animated mathematical models**>

Here is the board of disease 1 and we imagine symptom 1 occurs in 50% of all patients. Similarly we imagine that symptom 2 occurs equally in 50% of all patients. Now disease 2 we imagine likewise that 50% of all patients have symptom 1 and that 50% of all patients have symptom 2.

<**Card to camera**>

So these two symptoms could not possibly distinguish between these two diseases. True or false? False. They cannot distinguish between these two if they are independent, but if they are dependent then this situation could arise.

<**Card narrates over previous animated mathematical models**>

Disease 1. Disease 2. Disease 1. And these two symptoms alone could completely distinguish between these two diseases.

<**Card to camera**>

We've only dealt with two symptoms but obviously we could have three or four and many clinicians recognise such patterns and pass on their knowledge in their teaching. The great practical difficulty is to estimate these dependent probabilities which are at the basis of this pattern.

<**End credits**>