The Diagnosis of Jaundice by the Computation of Probabilities

R. P. KNILL-JONES, MSc, MB, MRCP, Research Fellow,
Hospital Health Services Research Unit, University
Department of Medicine, Western Infirmary, Glasgow

The purpose of this article is to show how the principles of Bayesian probability outlined by Professor Lindley on page 197 of this issue can be used for the diagnosis of patients with jaundice. My interest in the use of statistical techniques for the diagnosis of the jaundiced patient began in 1968. At that time, members of the Liver Unit at King's College Hospital found that, in spite of an increasing amount of information available about these patients, diagnosis was not becoming easier, as might have been expected. Indeed, some of the information was confusing and some of the investigations being used were expensive and of uncertain value.

A prospective study was started and information is now available on over 400 patients. The information collected (Table 1) consisted of defined

TABLE 1. Groups of tests collected prospectively

History	45
Physical signs	12
Laboratory results	34
Scintiscan	11
	102

questions from the history and routine laboratory tests, all of which can be available within 48 hours of a patient's admission to hospital. In addition, other investigations were noted, such as the hepatic scintiscan. Liver biopsy results were not included in the data. A final diagnosis was established on each of 309 patients by a combination of clinical outcome, result of liver biopsy, or result of surgery (Table 2). The information about these patients comprises a data base from which, in Professor Lindley's terminology, we can obtain the probabilities of a symptom being present given one of the diseases causing jaundice (p $C|A$).

TABLE 2. Data base of 309 cases. Percentage in each final diagnostic category

Medical cases 63·1%		Surgical cases 36·9%	
Acute viral hepatitis	18·1	Tumour of extra-hepatic biliary tree	25·6
Fulminant hepatic failure	8·1		
Cirrhosis	8·1	Gallstones	11·3
Active chronic hepatitis	8·1		
Primary biliary cirrhosis (PBC)	5·8		
Secondary tumour in liver	4·5		
Chlorpromazine jaundice	3·9		
Other drug jaundice	3·6		
Gilbert's syndrome	2·9		

This information can readily be used for the calculation of the probability that a new patient has one of the diseases, given that the patient has certain symptoms. The calculation is illustrated in Table 3, in which, for simplicity, only four diseases are used. The method readily extends to symptoms that may have more than just a 'yes' or 'no' answer. A modification of the data has to be used for continuous variables, such as age and serum bilirubin when applying Bayes's theorem (Bailey, 1965).

TABLE 3. Calculation of probabilities

	Symptom		p $(C\|A)$			p $(A\|C)$
Disease	Weight loss	Abdominal pain	Pruritus	Off Smoking		Probability
Drug jaundice	0·56	0·15	0·90	0·29	0·02	0·08
Hepatitis	0·66	0·43	0·62	0·68	0·12	0·5
PBC	0·70	0·09	0·78	0·25	0·01	0·04
Extra-hepatic tumour	0·78	0·37	0·72	0·45	0·09	0·38
					Sum = 0·24	1·0

For a patient with jaundice who has lost weight, has had abdominal pain, is itching and no longer enjoys smoking, the computation of probabilities proceeds for 'drug jaundice' thus—$0{\cdot}56 \times 0{\cdot}15 \times 0{\cdot}90 \times 0{\cdot}29 = 0{\cdot}02$, and similarly for the other diseases given in the table. Each of the numbers in the fifth column is divided by the sum of that column to give the probability that the patient described above has each one of the diseases. Thus, $0{\cdot}02 \div 0{\cdot}24 = 0{\cdot}08$, the probability of 'drug jaundice' for this patient. For simplicity, this table does not include the use of p (A), the prior probability of each disease.

The data base was tested on 65 new patients (Knill-Jones *et al.*, 1973) and used to predict whether each patient had 'surgical' jaundice or 'medical' jaundice. In 52 patients (80 per cent) a medical or surgical diagnosis was reached with a probability greater than 0·96, and at this point calculation stopped, corresponding to odds of at least 24 to 1 in favour of one of the decisions. Diagnoses reached with a probability of this level are referred to as

'certain' diagnoses, the remainder being referred to as 'uncertain' diagnoses. In these 52 patients an accuracy of 94 per cent was achieved; there were three errors (Table 4) as might be expected by using a probability of 0·96 as a stopping point. In the remaining 13 patients, 9 were correctly allocated, but with much lower probabilities. The method identifies this group of patients as needing more investigation and the computer program used for the calculations indicates which further estimations (such as scintiscan, or immunoglobulin) are most likely to lead to a 'certain' diagnosis.

TABLE 4. Diagnostic accuracy in 52 cases in which the probability of a medical or a surgical decision was greater than 0·96.

	Calculated		
True	Medical	Surgical	Total
Medical	37	2	39
Surgical	1	12	13
			52

Of the 39 medical cases, 2 were incorrectly calculated to have a surgical cause for jaundice. One patient had severe weight loss and prolonged jaundice, suggesting a pancreatic tumour, but biopsy suggested cholestatic viral hepatitis; the second had rigors suggesting gallstones but proved to have hepatitis.

One of the 13 surgical patients was regarded strictly as an error although no surgery could have helped her because a carcinoma of the gall bladder had already invaded the liver.

The choice of a particular probability for a stopping point depends on the kind of diagnosis being made. The higher the probability chosen, the more accurate the prediction in those cases reaching that probability (Knill-Jones *et al.*, 1973); for example, in distinguishing between acute hepatitis and the first episode of chronic active hepatitis within 48 hours of hospital admission, a much lower probability would be acceptable to many physicians. Once an agreed stopping point is reached the use of more information is most unlikely to alter a diagnosis. Further investigation in these cases would therefore be a wasteful use of resources that should be applied to the further investigation of those cases with an 'uncertain' diagnosis.

In the 52 'certain' cases, an average of 34 out of a maximum of 102 tests were used if each were selected at random. This suggests that, on average, up to two-thirds of the tests were redundant and illustrates Professor Lindley's

point that better handling of data often means reduction from large quantities to smaller quantities that are of real significance.

Further improvement occurs in these cases if sequential selection of tests is used. At each step in a sequential process a test is selected and the relative probabilities of possible diseases change according to the result for the particular patient. If each test is selected in such a way as to reduce uncertainty as quickly as possible, then the average number of tests needed for the 52 'certain' cases falls to 14. This reduction suggests that even more data is redundant than the two-thirds suggested above. The process is illustrated graphically on a probability diagram in Fig. 1 (*also see* Stokes, 1973). Further reduction is still possible (Teather, 1975).

Prospective testing of the data base on new cases arising in a teaching hospital is not the only experiment that must be done, since most cases of jaundice are seen in regional hospitals. The $p(C|A)$ will be unchanged if each disease behaves in the same way in regional as well as in teaching hospitals. There is evidence in other specialties that this is so even between hospitals in different countries (Jennett, 1975). However, the prior incidence of disease, $p(A)$, may well be different in the two types of hospital. In jaundice this may not be of great importance since no disease considered in the calculations is more than ten times more common than the rarest. This is also true in a different series of liver disease cases (Fraser and Franklin, 1974). The evidence from results of tests can easily swamp the evidence derived from the initial probabilities even when the rarer diseases are calculated to be the most likely diagnosis.

The study therefore continued in the teaching hospital and was extended to include patients in regional hospitals in South-East London. The data base was enlarged to include the 65 new cases mentioned above. At the present time there are 159 cases in the 'certain' group in the extended study. Diagnostic accuracy at King's College Hospital remains nearly the same at 95 per cent and is matched by 94 per cent accuracy in the regional hospitals. More cases in the regional hospitals have jaundice as a complication of other diseases (such as heart failure or septicaemia) and this will affect the choice of additional diseases for future studies (Stern *et al.*, 1975).

The calculation of one of 11 diagnoses is naturally more specific than calculation of either a 'medical' or 'surgical' diagnosis. Using the information available early in the patient's admission, a diagnostic accuracy of 71 per cent for the 65 cases in the original test group was achieved. In 55 per cent of these cases a 'certain' diagnosis was reached and the accuracy was 89 per cent. In the 'uncertain' cases a differential diagnosis is given, and this cannot easily be expressed as an overall accuracy figure. In the extended study there were

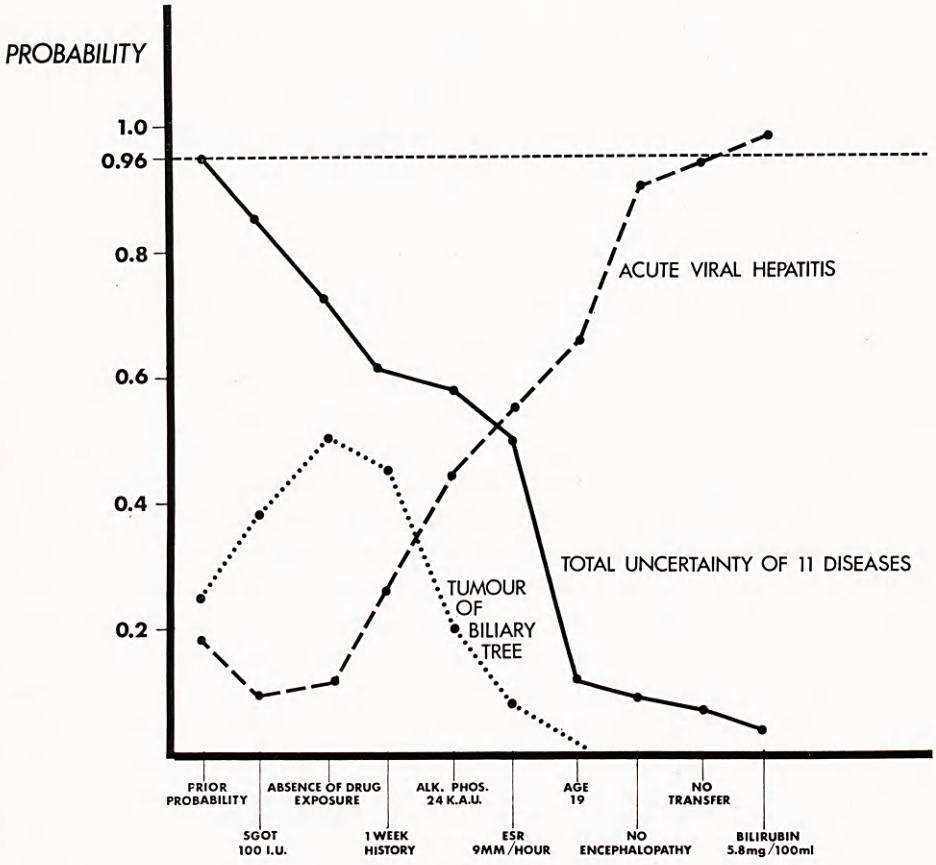

Fig. 1. Probability diagram showing the effect of test results from a patient in changing the probabilities for 2 of 11 diseases causing jaundice. For example, the modest elevation of alkaline phosphatase to 24 K.A.U. is more typical of hepatitis than of the higher values found in patients with pancreatic tumours. The probability of hepatitis thus rises and the probability of tumour falls when this test is used in the computation of probabilities. The final diagnosis is hepatitis.

Total uncertainty is the sum of the probability of each disease multiplied by its logarithm, and scaled on a range 0 to 1. Initial uncertainty is not at its maximum of 1 because the prior probabilities of diseases are not equal and therefore provide some information.

(Courtesy of the *British Medical Journal*.)

116 'certain' cases (53 per cent of the sample) and the regional cases among these were diagnosed with an accuracy of 88 per cent, comparable to the teaching hospital cases in the earlier study.

These results are encouraging but need to be compared to the accuracy of

clinicians. There are suggestions that an accuracy of 85 to 87 per cent is usually reached by clinicians faced with a jaundiced patient (Cattaneo *et al.* 1972; Martin *et al.*, 1960). A comparison of the accuracy of a computer and clinician is difficult to do, in jaundice, in such a way that the amount of information available to both is the same. In a preliminary study in which 20 case histories were abstracted and given to a group of clinicians, and identical information given to the computer, the clinicians' accuracy ranged from 45 to 65 per cent. The computer was correct in 70 per cent; in two cases the true diagnosis was not one of the 11 diseases and its accuracy could not therefore have been higher than 90 per cent (Stern *et al.*, 1974).

What are the implications of the use of probability calculations in the management of jaundice? The method achieves similar or possibly greater accuracy than clinicians without extensive knowledge of liver disease, by the use of the cheapest and most readily available tests. This stage can be done in the outpatient department. It identifies a group of cases in which further investigation is necessary and can suggest the most appropriate investigation. The early identification of a surgical group of patients promises some reduction in the use of hospital beds for the pre-operative investigation of these patients.

Acknowledgement

This work was supported in part by the Nuffield Provincial Hospitals Trust.

References

Bailey, N. T. J. (1965) *Proceedings of Medical Research Council Conference on Mathematics and Computer Science in Biology*, No. 103. London: H.M.S.O.

Cattaneo, A. D., Luchelli, P. E., Rocca, E., Mattioli, F. and Becchi, G. (1972) *Abdominal Surgery*, **14**, 71.

Fraser, P. M. and Franklin, D. A. (1974) *Quarterly Journal of Medicine*, **169**, 73.

Jennett, B. (1975) *Journal of the Royal College of Physicians of London*, **9**, 231.

Knill-Jones, R. P., Stern, R. B., Girmes, D. H., Maxwell, J. D., Thompson, R. P. H., Williams, R. (1973) *British Medical Journal*, **1**, 530.

Martin, W. B., Apostalakos, P. C., and Roazon, H. (1966) *American Journal of Medical Science*, **240**, 571.

Stern, R. B., Knill-Jones, R. P., and Williams, R. (1974) *Methods of Information in Medicine*, **13**, 79.

Stern, R. B., Knill-Jones, R. P., and Williams, R. (1975) Paper in preparation.

Stokes, J. F. (1973) *Journal of the Royal College of Physicians of London*, **8**, 5.

Teather, D. (1975) *Journal of the Royal College of Physicians of London*, **9**, 219.