

1976

A statistical model for the diagnosis of neonatal heart disease

Dennis Marshall Fisher
Yale University

Follow this and additional works at: <http://elischolar.library.yale.edu/ymtdl>

Recommended Citation

Fisher, Dennis Marshall, "A statistical model for the diagnosis of neonatal heart disease" (1976). *Yale Medicine Thesis Digital Library*. 2586.

<http://elischolar.library.yale.edu/ymtdl/2586>

This Open Access Thesis is brought to you for free and open access by the School of Medicine at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Medicine Thesis Digital Library by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

YALE MEDICAL LIBRARY



3 9002 08627 8703

A STATISTICAL MODEL FOR
THE DIAGNOSIS OF NEONATAL HEART DISEASE



DENNIS MARSHALL FISHER

1976

YALE




MEDICAL LIBRARY

YALE



MEDICAL LIBRARY



Digitized by the Internet Archive
in 2017 with funding from
The National Endowment for the Humanities and the Arcadia Fund

<https://archive.org/details/statisticalmodel00fish>

A STATISTICAL MODEL FOR
THE DIAGNOSIS OF NEONATAL HEART DISEASE

A thesis submitted to the faculty of
Pediatrics, Yale University School of
Medicine, in partial fulfillment of the
requirements for the degree of Doctor
of Medicine.

DENNIS MARSHALL FISHER

MARCH, 1976

NEW HAVEN, CONNECTICUT

Permission for photocopying or microfilming of "A STATISTICAL
MODEL FOR THE DIAGNOSIS OF NEONATAL HEART DISEASE" "

(TITLE OF THESIS)

for the purpose of individual scholarly consultation or reference is hereby granted by the author. This permission is not to be interpreted as affecting publication of this work or otherwise placing it in the public domain, and the author reserves all rights of ownership guaranteed under common law protection of unpublished manuscripts.



Signature of Author Dennis M. Fisher

4/14/76
Date

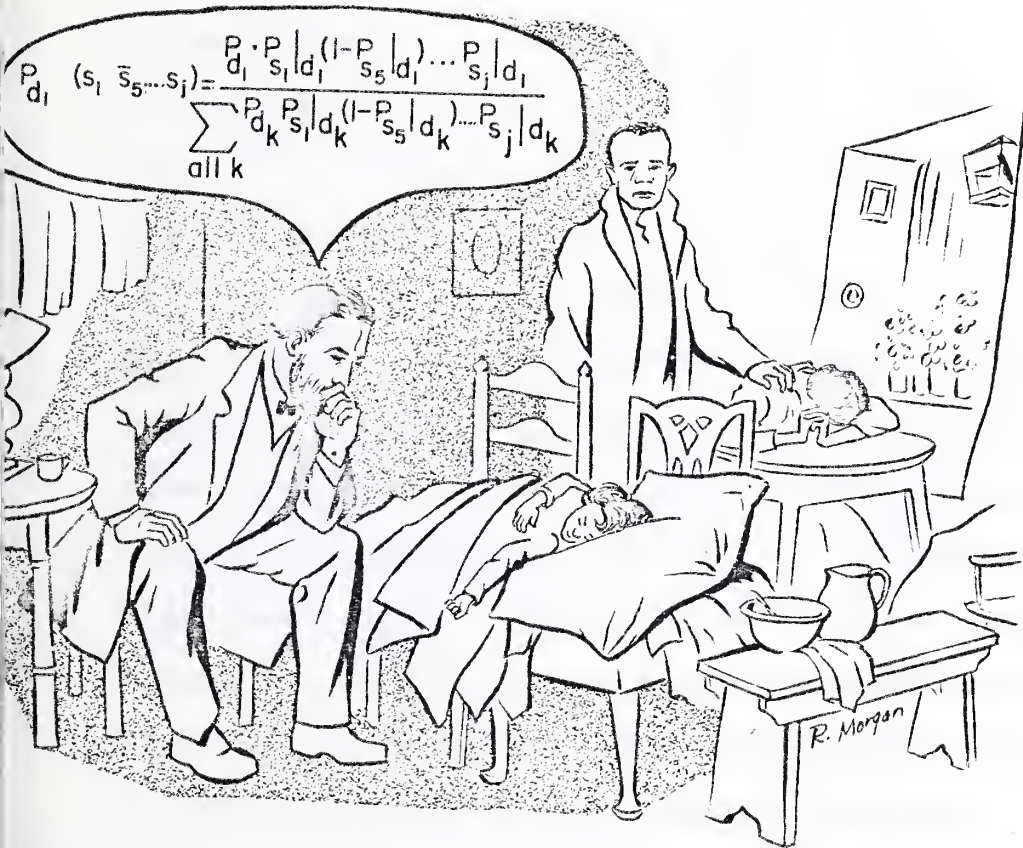


FIG. 1. Bayes' theorem in practice.

From: Lodwick, Gwilym, et al., "Computer Aided Diagnosis of Radiographic Images" Journal of Chronic Diseases (Volume 19, No. 4, April, 1966) p. 486.

ABSTRACT

During the last three decades, numerous authors have attempted to simulate the process through which clinicians make diagnoses. Limited originally to simple mechanical models^{1,2}, the advent of electronic computers created new technologies³. With the introduction of Boolean Algebra⁴, it became possible to apply Bayes Law (a formula which allows predictions of disease probabilities from incidence data and the occurrence of symptoms within diseases) to clinical problems.

Studies in the radiographic diagnosis of primary bone tumors⁵, the acute abdomen⁶, thyroid disease⁷ and Cushing's syndrome⁸ have demonstrated the value of Bayesian models. Warner et al.⁹ studied congenital heart disease in a population ranging from one month to greater than twenty years. Their work omitted the neonatal population, the age during which the most serious cardiac malformations manifest. This study applies the technique of Bayesian analysis to a neonatal population. The model involves identification of the important symptoms (historical, physical findings, EKG, roentgenographic findings) and diseases followed by completion of a symptom-disease matrix (the incidence of each disease as well as the incidence of each symptom within each disease). Using a computer to tabulate the enormous quantities of data, the program's accuracy ranges from fifty-two to sixty percent in different

populations, compared to an approximate clinical accuracy of eighty percent. With variations in the symptom-disease matrix to correct errors in the data, the accuracy of the computer program can be increased to seventy-five percent. In addition, the computer compiles a differential which not only resembles the clinical differential but at times is considered more complete or logical.

With improvements in the data base, it is expected that the program's accuracy can be greatly increased, and could become an effective adjunct to clinical diagnosis. Other studies have demonstrated the utility of Bayesian models as clinical¹⁰ and teaching aids¹¹. The possibilities, including simulation of complex clinical problems^{12,13}, are considerable. As more models are developed, the limiting factor will be physicians' acceptance of computer-based statistical models.

FOOTNOTES

1. F. A. Nash, "Differential Diagnosis" The Lancet (Volume 22, No. 6816, April 17, 1954) pp. 874-875.
2. Martin Lipkin, James Hardy, "Differential Diagnosis of Hematological Diseases Aided by Mechanical Correlation of Data" Science (Volume 125, March 22, 1957) pp. 551-552.
3. Lewis Barnes, et al., "Computer-Assisted Diagnosis in Pediatrics" American Journal of Diseases of Children (Volume 127, June, 1974) pp. 852-858.
4. Robert Ledley, Lee Lusted "Reasoning Foundations of Medical Diagnosis" Science (Volume 130, No. 3366, July 3, 1959) pp. 9-21.
5. Gwilym Lodwick, et al., "Computer Diagnosis of Primary Bone Tumors" Radiology (Volume 80, February, 1963) pp. 273-375.
6. J. C. Horrocks, et al., "Computer-Aided Diagnosis" British Medical Journal (April 1, 1972) pp. 5-9.
7. Lawrence Fitzgerald, et al., "A Computer Program for the Diagnosis of Thyroid Disease" American Journal of Roentgenology (Volume 97, No. 4, August, 1966) pp. 901-905.
8. C.A. Nugent, et al., "Probability Theory in the Diagnosis of Cushing's Syndrome" Journal of Clinical Endocrinology (Volume 24, No. 7, 1964) pp. 621-627.
9. Homer Warner, et al., "A Mathematical Approach to Medical Diagnosis" Journal of the AMA (Volume 177, No. 3, July 22, 1961) pp. 558-567.
10. E. T. deDombal, "Surgical Diagnosis Assisted by a Computer" Proceedings of the Royal Society of London (Volume 184, 1973) pp. 443-440.
11. J. C. Horrocks, et al., "Production of Artificial Case Histories" British Medical Journal (June 5, 1971) pp. 578-581.
12. Anthony Gorry, et al., "Sequential Diagnosis by Computer" Journal of the AMA (Volume 205, No. 12, September 16, 1968) pp. 849-854.
13. A. S. Ginsberg, Decision Analysis in Clinical Patient Management with an Application to the Pleural Effusion Syndrome (Santa Monica, Rand Corporation, 1971)

CONTENTS

CHAPTER I		Page	1
CHAPTER II	The Literature	Page	6
CHAPTER III	Bayes Theorem	Page	49
CHAPTER IV	Methodology	Page	62
CHAPTER V	Results and Discussion	Page	88
CHAPTER VI	Beyond Bayes	Page	106

	Bibliography	Page	114
--	--------------	------	-----

APPENDIX 1	Data Collection Form		
APPENDIX 2	Computer Program		
APPENDIX 3	Symptom-Disease Matrix		
APPENDIX 4	Sample Patients		
APPENDIX 5	Sample Patients - Reanalyzed with Incidence Data Omitted		

CHAPTER I

The process by which the practitioner of medicine formulates clinical decisions is complicated. In its simplest form, it can be seen as having three components. First, the clinician interacts with the patient to make observations and accumulate data. This is in the form of obtaining a history, performing a physical examination and ordering the appropriate laboratory procedures. Second, the clinician takes the information he has gathered and translates it into a diagnosis, an assessment of the disease process with which the patient presents. Finally, the clinician formulates a plan of therapy which reflects both the illness and the opportunities for intervention through medical care, social or possibly economic change.

To aid the clinician in learning the first of these tasks, there are many resources. Early in medical training, he is exposed to a course in physical diagnosis. Numerous textbooks offer the trainee techniques of examination. During clinical training, through constant exposure and experimentation, the student is able to secure an adequate knowledge of physical diagnosis.

To formulate a plan of therapy, the clinician is again offered a large number of resources. Medical

literature abounds with textbooks and journals, and faculty and housestaff are available to help with clinical problems. Having identified the disease under consideration, the physician need only turn to these sources to obtain information about a clinical problem.

It is the middle component - the translation of data into diagnosis - that challenges the trainee. Often, the diagnosis is readily apparent from the history, physical findings or laboratory data. A student, examining a patient with the recent onset of fever, sputum-productive cough and an infiltrate on chest roentgenogram, need not have great insight to suspect the presence of a pneumonia. Frequently, however, the diagnosis is more complicated, either from the difficulty of obtaining sufficient historical, physical examination or laboratory data, variation of the patient from the classical case or the spectrum of presentations of the same disease.

Senior clinicians, when asked to explain their diagnostic reasoning, may employ the term "clinical judgment". This implies, "I recognize something in this patient which clues me towards a specific diagnosis". In identifying their decision process, some invoke the presence of an obscure finding, others may describe a patient with a similar presentation they saw at some other time in their career. Astute clinicians are often

able to outline their decision processes, making explicit the manner in which they accumulate and process data. An anecdote offered by Mack Lipkin, a psychiatrist with an interest in early medical training, describes an interesting example:

When first seen, the patient aged seventy, was lying on his back in a dimly lit room literally screaming because of violent abdominal pain. Only two things were instantly apparent to the other physician and me - enormous abdominal distension and extreme generalized abdominal tenderness. The history and further examination showed that the patient had peritonitis of unknown origin. He had been ill for four days but had refused to call a physician. He did not remember where the pain began or whether he had had a bowel movement. When the surgeon arrived, he was given the information: four days of pain, obvious peritonitis, and a temperature of 105 degrees Fahrenheit. He walked into the patient's bedroom, took one look, and in about five seconds, said, "Ruptured retrocecal appendix." Asked why he was so sure that the offending organ was the appendix, that it was ruptured, and that it was retrocecal, he looked surprised. He said, "What else can it be? Look where he has burned himself with his hot water bottle." The other physician and I looked at the abdomen which had the reticulated pattern that results from continued use of the heating pad, stared at each other, and then turned to the surgeon. "Why do heating pad burns mean a ruptured retrocecal appendix?" He answered, "But look where the marks are darkest." We looked carefully and could see no difference. He pointed to the extreme right side of the abdomen. We looked again and agreed that perhaps these marks were slightly darker. He explained that he knew of no other condition that would start with pain in that part of the extreme right side of the abdomen and go on to produce peritonitis except a ruptured retrocecal appendix. His diagnosis was absolutely correct.

Rarely is the clinician called upon to make such difficult observations. Nonetheless, he often makes diagnoses without having made specific observations on all aspects

of the patient or making explicit statements regarding his observations.

During the last three decades, the process by which clinicians make diagnoses has come under careful scrutiny. Initially, theoreticians attempted to build simple mechanical models to aid in the teaching of certain diagnostic problems. In the nineteen fifties, Boolean algebra and logic, with their ability to translate language into mathematical or logical statements, offered researchers the opportunity to simulate simple clinical models. With the development of high speed electronic computers with their massive storage capacities, some authors have simulated more complex clinical situations.

The purpose of this dissertation is three-fold. First the author will review the history of these diagnostic models. Then, a clinical model which has been developed and tested for use in the diagnosis of congenital heart disease in the newborn period will be presented. Finally, there will be a discussion of the implications of these diagnostic models in present-day clinical medicine, as well as some of the opportunities for future research and clinical application.

FOOTNOTES

1. Mack Lipkin, The Case of Patients (New York, Oxford University Press, 1974) pp. 155-156.

CHAPTER II

THE LITERATURE

Confronted with each new medical problem, the physician begins by recalling recognized patterns of disease. To systematize his thinking, although not necessarily consciously, he calls into action an algorithm, a procedure for solving a problem.

The simplest algorithm - one symptom or some combination of symptoms or signs stimulating the physician towards a particular diagnosis or category of diagnoses - can be represented symbolically as:

Symptom(s) —————> Disease(s)

To further this process, the clinician may call upon a textbook of differential diagnosis, a resource structured similarly to the algorithm shown above. Considering each of the symptoms separately, the physician compiles lists which may overlap partially, completely, or not at all. Extracting the common elements, he is often able to make a significantly smaller list, and if additional data is made available, the list of diagnostic possibilities can be limited even further. Through this process, the clinician has extended the linear model shown above to a more complex one, a model

which involves the intersection of different sets.

A Mechanical Model

If the number of symptoms is extensive, the textbook of differential diagnosis comprehensive, or more than one disease present, the clinician may have created an enormous task, the culling and sorting of hundreds of different diseases before the appropriate diagnosis can be entertained. To simplify this task, Nash¹ outlined in 1954 a tool which he felt could be beneficial for medical diagnosis. The device, which viewed on side resembles a large slide rule, consists of a frame into which a number of different strips can be inserted (Figure 1). Into one edge of the frame, disease names are inscribed (three hundred in this case, but limited only by space). The strips, each of which represents data obtained from history, physical examination or laboratory procedures, are slipped into the frame, one beside the other. Markings on the "symptom" strips then line up adjacent to the diseases in which they are present. By selecting strips to represent the patient's symptoms and observing the frame for the location of the greatest number of markers, the clinician is able to simplify his differential diagnosis.

While rather cumbersome, this device is an attempt to simplify the use of textbooks of differential diagnosis.

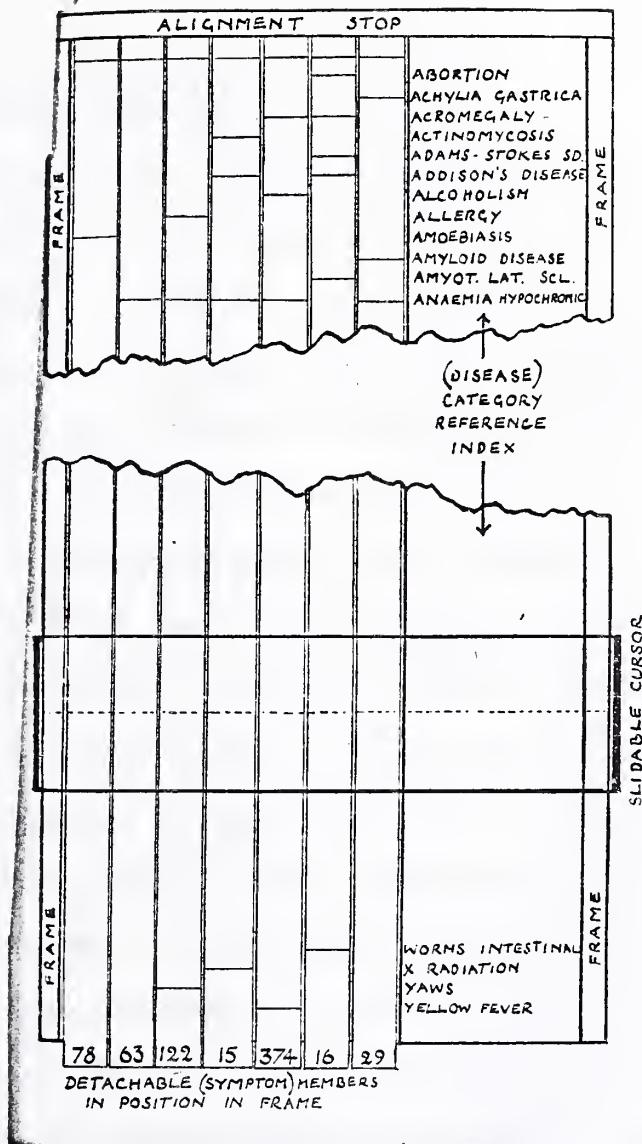


FIGURE 1

A SIMPLE MECHANICAL MODEL

From: Nash, F. A., "Differential Diagnosis" The Lancet (Volume 22, No. 6816, April 17, 1954) p. 875.

Proposed before the era of computational devices, it applied the technology of the era.

Early Data Punching

Three years later, as punched cards became available to the scientific community, Lipkin and Hardy² utilized the newer technology. They proposed to "evaluate the efficiency with which mechanical classification and correlation of data might assist in the utilization of data in the differential diagnosis of ... disease."³ Using marginal punched cards (Figure 2), a type of data card in which each hole punched around the edge of the card represents a specific piece of information, one space corresponding to every observation to be considered, they selected 138 findings from history, physical examination, peripheral blood, bone marrow and other laboratory examinations. Twenty-seven hematological diseases were classified according to these criteria and a "master card" prepared for each disease by punching a wedge between the appropriate hole and card edge.

Eighty case records from a university hospital were then abstracted and the appropriate information punched on marginal cards. The cards were tabulated, apparently by comparing each patient card to each of the 27 disease cards. For fifty of the cases an exact correlation to one disease was obtained, and in each it was identical

to the clinical diagnosis obtained from the hospital chart. In twenty-three, the data from the case corresponded to more than one of the master cards. For these, additional data was obtained from the charts and when retabulated each of the diagnoses was made precisely. In only seven cases was the data not identical to any of the master cards. In returning to these medical records, each patient was found to have more than one hematological disease.

A year later, Lipkin and Hardy⁴ presented a more sophisticated model using the same data in conjunction with mechanical techniques. Now using 26 diseases punched as before, they took advantage of the opportunities offered by marginal punching, a process termed "direct sorting"⁵. To find all the diseases characterized by a single observation, for example, splenomegaly, "one would place this set of cards front to back and place a metal or plastic rod into the hole to which that item had been assigned. When the rod was raised, cards representing diseases characterized by a large spleen would fall, because the triangular wedge would have been punched into that space. Cards without the wedge would be raised"⁶. Discarding the eliminated diagnoses and inserting a rod corresponding to a second symptom, the number of "eligible" diagnoses could be further limited. When the symptom list was exhausted several possibilities could occur:

no disease, only one, or more than one remained in the differential. As before, in 73 of the 80 cases a single correct diagnosis was obtained. The elimination of all possible diagnoses never occurred while in seven cases, additional mathematical manipulation was necessary to obtain the diagnosis.

Electronic Models

With the development of high-speed electronic computers, attempts were made to utilize the newer technologies. Barness⁷ proposed a computer-assisted diagnostic program which presently include 1,500 different diagnoses. The patient's symptoms, selected from a comprehensive list of 650 abnormal findings (75 historical, 315 physical, 60 roentgenographic, and 200 laboratory findings) are entered via a teletypewriter to a remote computer facility. "The computer systematically searches its data bank, comparing the patient's findings with each diagnosis. The computer selects the diagnoses that should be included in a complete differential diagnosis. These are listed in a computer report by the teletypewriter terminal"⁸.

To apply the system to clinical situations hospital records were reviewed to find pediatric patients whose final diagnosis was not included in the admitting differential. Twenty charts were abstracted and abnormal

findings selected from the initial history. When tabulated by the program, for eighteen of the twenty cases the differential provided by the computer included the discharge diagnosis. Although the program is presently extremely expensive - each run costs \$50 - the authors contend that with general use its price could become "similar with that of a customary laboratory test"⁹. In addition, the differential diagnosis is often too long to be clinically useful.

Decision Trees

Alternatively, one can employ decision trees as diagnostic aids. A typical one, proposed by Krovetz¹⁰ (Figure 3) offers the user a number of decision nodes from which to choose. After each selection is made, he reaches another node, until finally, a minimal number of diagnoses can satisfy the acknowledged criteria.

Decision trees, while very useful in simple and "classical" cases, tend to have significant limitations. Atypical cases, or those which present in varied ways, tend to become lost in the decision process. For example, Krovetz's cardiological model uses "obvious cyanosis" as its first nodal point. Following each of the superior alternatives - no cyanosis, no thrill, decreased pulmonary markings - one reaches a point at which no diagnoses are tenable. The authors then suggest reexamination for

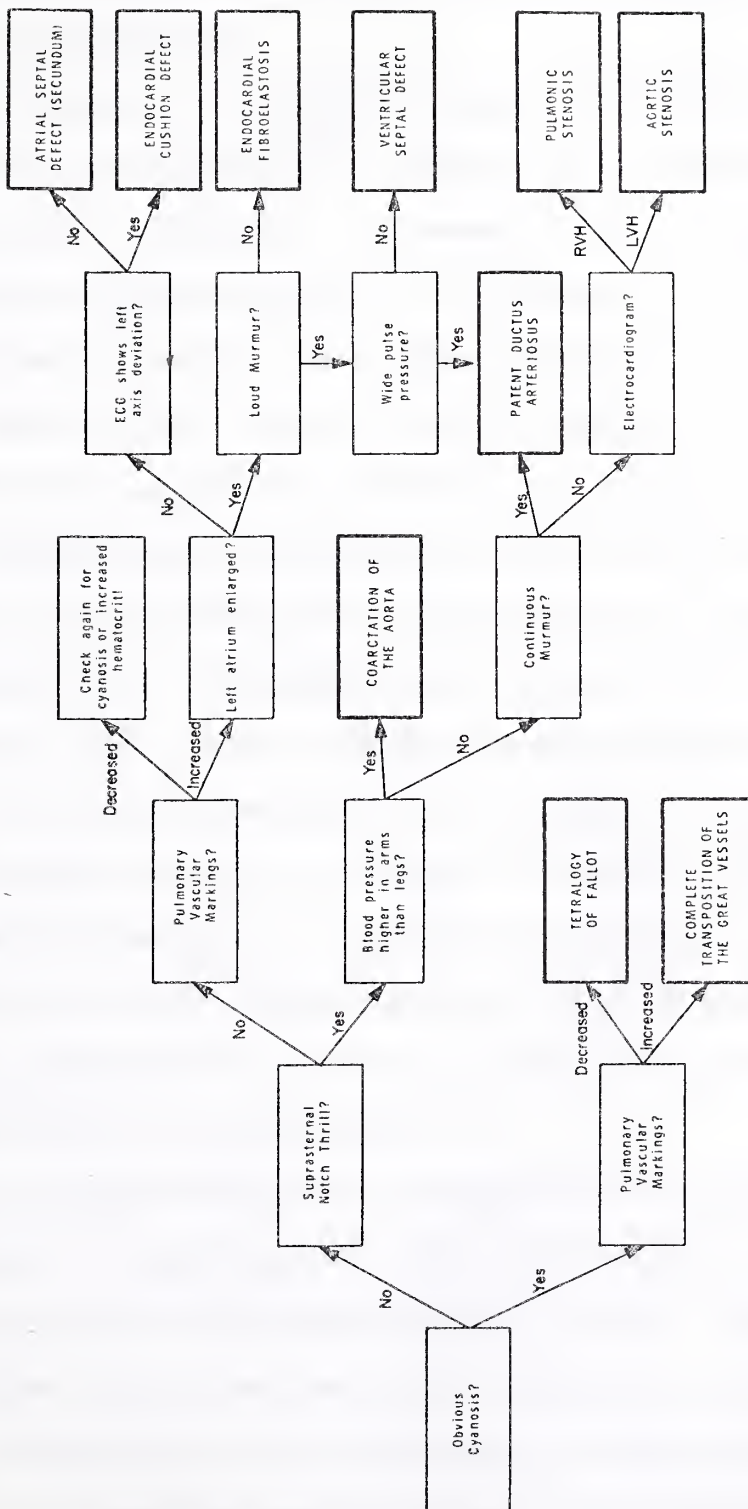


Figure 211. Schema of differential diagnoses after one year of age.

FIGURE 3

A branching sequence for the diagnosis of congenital heart disease.

From: Krovetz, Jerome, et al., Handbook of Pediatric Cardiology (New York, Harper and Row, 1969), p. 367.

cyanosis, allowing for either observer error or variations in hematocrit.

An attempt to circumvent these limitations was proposed by Kleinmuntz¹¹. His decision tree involves neurological diagnosis, "an example par excellence of diagnostic problem solving ... because of the highly structured nature of the clinical data within it."¹² The scheme evolves from a game of "Twenty Questions" in which the experimenter thinks of a disease while the other participants are invited to ask questions in pursuit of the diagnosis. The information provided by the neurologist, characterizing symptoms S1,S2,S3,etc. (Figure 4) may involve history, physical exam or laboratory data. Following completion of the game, a diagnostic tree is built, which is intended "to be an information processing theory of ... diagnostic behavior.... [The] program is a theory that explains the diagnostic problem solving behavior of individual clinical neurologists in a given set of n situations."¹³

Two distinct models of diagnosis are proposed by Kleinmuntz. The computer could "have stored in its memory core all the known symptoms, signs, demographic variables, laboratory tests and treatments, and then when a particular set of symptoms or signs are presented to it, the computer could proceed very much like a beginning medical student who might consult a diagnostic

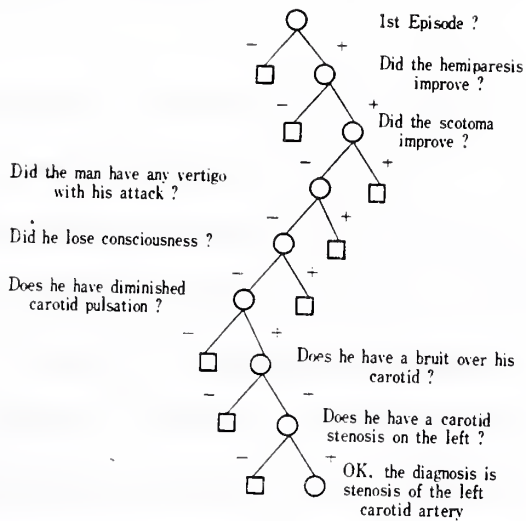


FIG. 1. A tree structure of a neurologist's diagnostic game in which the information given was: Sudden left central scotoma and right hemiparesis in a 55 year old.

FIGURE 4

A neurologist's branching sequence.

From: Kleinmuntz, Benjamin, "Diagnostic Problem Solving by Computer" Japanese Psychological Research (Volume 7, No. 4, 1965) p. 192.

manual or textbook."¹⁴ In addition to being very tedious, this program would teach very little about the "diagnostic process". Alternatively, the program could incorporate all the above information "and take into consideration the relative frequency of a set of symptoms for a particular disease and the frequency of occurrence of a particular disease in a particular geographic region."¹⁵

One model employing these techniques appears in a recent paper by Brand, Dove and Meyers.¹⁶ To characterize the difference between non-bacterial and bacterial meningitis, they examined the records of several hundred children seen at the Yale New Haven Hospital during a fifteen year period. Several innovations appear in their study including derivation of "cutoff points" from data analysis rather than arbitrary criteria, and use of statistical techniques rather than "clinical judgment" to determine which criteria discriminate "best". Above all, the branching sequence does not offer only one diagnosis at the completion of all the nodes but considers at all points the likelihood of bacterial versus non-bacterial meningitis (Figure 5). Only if the likelihood were zero or 100% would this be equivalent to the traditional branching sequence.

COMPLETE PROBABILITY TREE
 DISTRIBUTION OF CASES BASED ON COMBINED PATIENT AGE, CSF GRAM STAIN, CSF GLUCOSE LEVEL
 AND CSF WHITE CELL COUNT

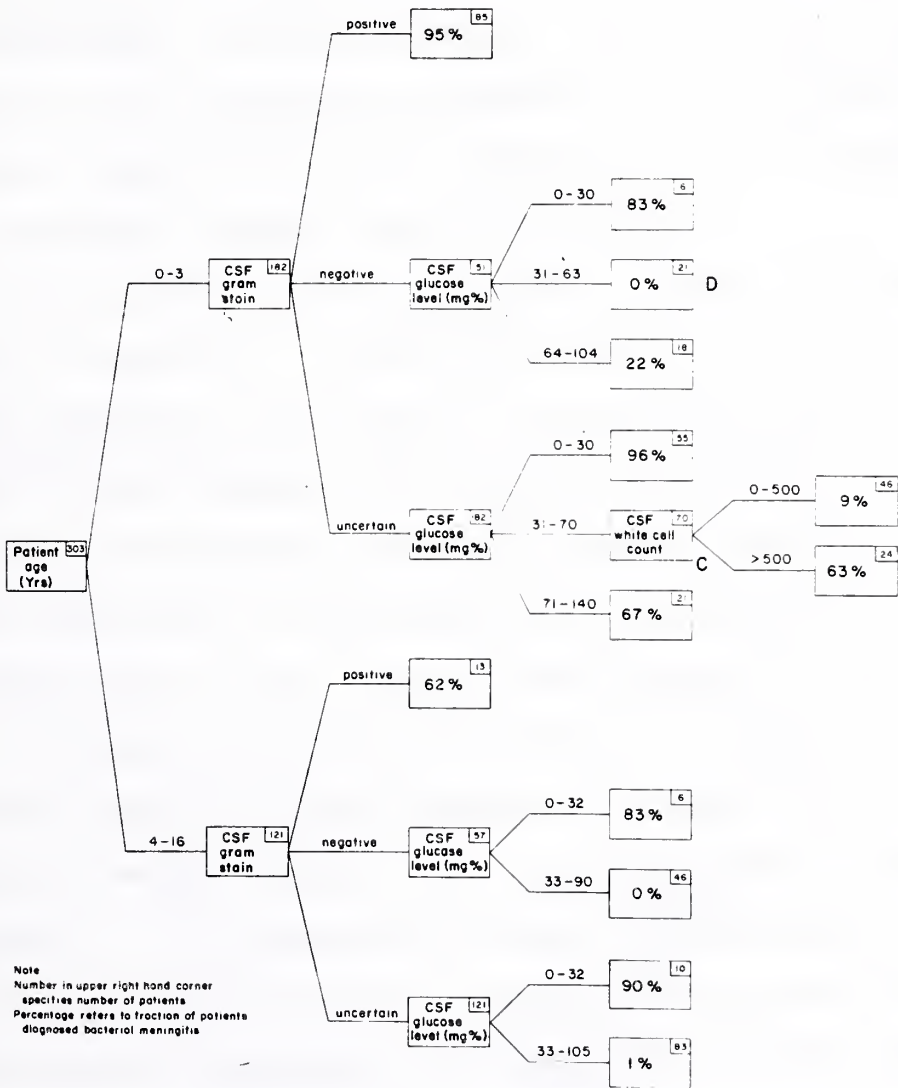


FIGURE 5

A branching sequence for the differentiation of non-bacterial and bacterial meningitis.

From: Brand, Don, et al., "A Technique for Analyzing Clinical Data to Provide Patient Management Guidelines: A Study of Meningitis in Children" Submitted for publication.

Statistical Models

In 1964, Lipkin¹⁷ proposed a statistical model to supplement his earlier work. Recognizing that the marginal card technique was unable to differentiate between important and less important data, he sought to weigh the relative value of symptoms.

Using twenty-six hematological diseases, he listed the 138 symptoms. "A numerical value or weight was then assigned to each item of information present in each of the diseases. The instructions stated: given a disease description, assign a positive weight to data that contribute to the diagnosis, give a negative weight to data that do not. Let the significance of the item in the disease determine the weight.... The sum of all the weights in each disease will yield the total positive weight and the total negative weight of the disease.... By comparing the sum of the weighted diagnostic criteria of the hospital case to the sum of the weighted diagnostic criteria of the disease, the degree of identical fit of hospital case data in each case was defined."¹⁸

Figure 6 shows the profile of one disease, followed by the same profile superimposed on a patient profile, yielding a Diagnostic Score. The Diagnostic Index is:

$$D.I. = \Sigma X \cdot W / \Sigma W$$

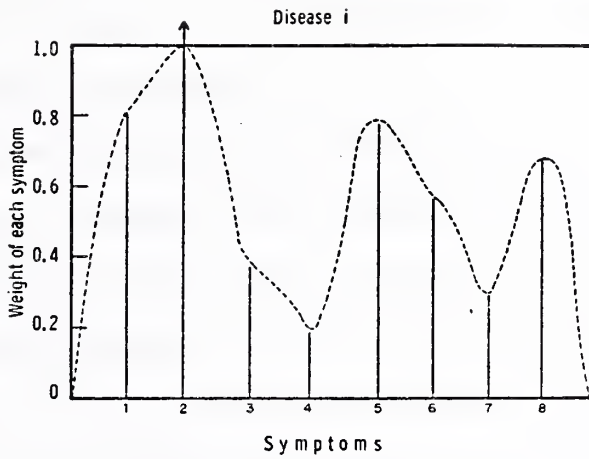


FIG. 2.—Schematic profile of one disease. The abscissa records symptoms 1-8; the ordinate records the weight of each symptom. The dotted line connects the weights and emphasizes groups of symptoms having high and low weights. The arrow defines a pathognomonic symptom.

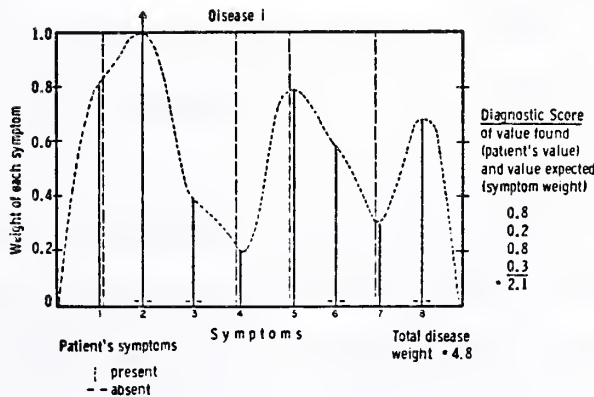


FIG. 3.—The same disease profile (solid lines) and the profile of a patient (dashed lines) superimposed. The value found (patient's value) is multiplied by the value expected (symptom weight) to produce a Diagnostic Score. Further division by the total disease weight produces the Diagnostic Index. The latter measures the degree of identical fit of the patient's data in each disease.

FIGURE 6

From: Lipkin, Martin, "The Likelihood Concept of Differential Diagnosis" Perspectives in Medicine and Biology (Summer, 1964), p. 488-489.

where X (1 or 0) corresponds to the symptom being present or absent. W is the weight assigned to that item in that disease.

"If a hospital case contained all the findings of a given disease, the Diagnostic Index was 1.00. If a hospital case contained none of the findings that characterized a disease, the Diagnostic Index was zero. If it contained one-half the findings of the disease, the Diagnostic Index was 0.5."¹⁹

Figure 7 demonstrates several representative cases, comparing hospital diagnoses to diagnostic indices. Lipkin notes that "when the Diagnostic Index approached 0.2 the disease in question became important in the differential diagnosis and had many features identical to those in the hospital case. When the Diagnostic Index reached 0.4, the disease was very important".²⁰

A second statistical model, for the diagnosis of thyrotoxicosis, was authored by Crooks, Murray and Wayne²¹ in 1958. Their method "consisted in allocating a positive or negative score to each clinical feature, the values being based on an analysis of the relative frequency of symptoms and signs in the disease. In this way a total score, or clinical diagnostic index, can be obtained in each case."²² Signs and symptoms

COMPARISON OF WEIGHTED AVERAGES AND HOSPITAL DIAGNOSES
FOR SEVEN CASES

<i>Case</i>	<i>Diagnostic Index (weighted averages, positive)</i>	<i>Hospital diagnoses</i>
1	Chronic lymphatic leukemia 0.41 Acute hemolytic anemia 0.28 Idiopathic thrombocytopenic purpura 0.27 Hemophilia 0.26	Chronic lymphatic leukemia Acquired hemolytic anemia
2	Agranulocytosis 0.49 Chronic lymphatic leukemia 0.45 Sprue 0.21 Hereditary spherocytosis 0.19	Chronic lymphatic leukemia Agranulocytosis
3	Mediterranean anemia 0.28 Hemophilia 0.23 Acute hemolytic anemia 0.22 Multiple myeloma 0.22	Mediterranean anemia Nonthrombocytopenic purpura
4	Hypochromic microcytic anemia 0.31 Idiopathic thrombocytopenic purpura 0.27 Multiple myeloma 0.19 Mediterranean anemia 0.15	Idiopathic thrombocytopenic purpura Anemia, acute and chronic
5	Multiple myeloma 0.37 Aplastic anemia 0.34 Agranulocytosis 0.30 Chronic lymphatic leukemia 0.09	Multiple myeloma Agranulocytosis
6	Agranulocytosis 0.47 Chronic lymphatic leukemia 0.21 Eosinophilic leukemia 0.19 Monocytic leukemia 0.18	Agranulocytosis Macrocytic anemia
7	Agranulocytosis 0.47 Aplastic anemia 0.21 Chronic lymphocytic leukemia 0.19 Acute posthemorrhagic anemia 0.19	Macrocytic anemia Neutropenia Thrombocytopenia

FIGURE 7

Diagnostic indices for sample hematological cases

From: Lipkin, Martin, "The Likelihood Concept of Differential Diagnosis" Perspectives in Medicine and Biology (Summer, 1964), p. 487.

were chosen "because they had previously been shown by a clinical survey to differ in their incidence in thyrotoxic and normal subjects."²³ Weights were assigned to the signs such that "the positive or negative values of these scores were ... allocated on the basis of the relative diagnostic significance of each"²⁴ as found in previous studies (Figure 8). After some modifications necessary to accomodate for observer variation, the clinical diagnostic indices were calculated. In a group of patients who were clinically "non-toxic", the scores ranged from -16 to +10 while definitely "toxic" patients ranged from +21 to +42. One hundred eighteen patients, for whom an initial clinical diagnosis was unclear despite radioactive iodine studies, were placed in a "doubtful" group. Of the sixty-seven felt to be non-toxic by final clinical diagnosis, fifty-nine had diagnostic indices in the non-toxic range (less than 11), seven in the intermediate range and one in the toxic range (twenty or greater). Of these eventually felt to be toxic, none fell into the non-toxic range, six in the intermediate area and forty-five had diagnostic indices which correlated with their clinical status. Excluding the single non-toxic patient with a toxic index, these data would suggest that a diagnostic

TABLE I

Weighting Factors Allocated to the Symptoms and Signs of Thyrotoxicosis

<i>Symptoms of recent onset or increased severity</i>			<i>Signs</i>		
	<i>Present (score)</i>	<i>Absent (score)</i>		<i>Present (score)</i>	<i>Absent (score)</i>
Dyspnoea on effort	+1		Palpable thyroid	+3	-3
Palpitations	+2		Bruit over thyroid	+2	-2
Tiredness	+2		Exophthalmos	+2	
Preference for heat (irrespective of duration)		-5	Lid retraction	+2	
Preference for cold	+5		Lid lag	+1	
Indifferent to temperature	0		Hyperkinetic movements	+4	-2
Excessive sweating	+3		Fine finger tremor	+1	
Nervousness	+2		Hands:		
Appetite increased	+3		Hot	+2	-2
Appetite decreased		-3	Moist	+1	-1
Weight increased		-3	Casual pulse rate:		
Weight decreased	+3		Auricular fibrillation	+4	
			Regular rate:		
			Under 80		-3
			80-90	0	
			Over 90	+3	

FIGURE 8

From: Crooks, J., et al., "Statistical Methods Applied to the Clinical Diagnosis of Thyrotoxicosis" Quarterly Journal of Medicine (Volume 28, No. 110, April, 1959) p. 212.

index of ten or less represented a non-toxic patient while at greater than twenty, hyperthyroidism was present. In the intermediate range, a relatively small group, no conclusions could be drawn.

A third statistical model was proposed by Collen, et al.²⁵, for use in the Kaiser Foundation Health Plan. As part of a multiphasic screening program, an attempt was made to identify "whether there is sufficient likelihood of disease being present to warrant further specific diagnostic testing".²⁶ One example, the differentiation of asthmatic patients from those without asthma, involved a "likelihood" principle proposed by Neyman²⁷. The likelihood ratio "is the ratio of the probability (P_D^S) with which a selected set of symptoms (S) occurs in a specific disease (D) to the probability (P_N^S) with which the same set of symptoms (S) occurs in the non-diseased state (N); that is $\theta = P_D^S/P_N^S$. Each symptom set, therefore, is associated with a θ value. The θ s are then arranged in order of increasing magnitude."²⁸

"Six dichotomous questions were selected from a questionnaire form given to 230 patients with a clinical diagnosis of bronchial asthma and to a group of 517 randomly selected patients who ... were known to be free of asthma."²⁹ The results, shown in Figure 9 demonstrate an area in which the likelihood is zero and another with a likelihood of infinity. Interestingly, seven percent

Table 2—Proportion of Patients with (P_D^S) and without (P_N^S) Bronchial Asthma Answering Yes to Six Diagnostic Questions

Question Number						Asthma Patients		Nonasthma Patients		Likelihood Ratio	
1	2	3	4	5	6	No.	(P_D^S)	No.	(P_N^S)	$\theta = (P_D^S / P_N^S)$	
No	No	No	Yes	No	No	0	0	16	0.03095	0	
No	No	No	No	Yes	No	0	0	6	0.01161	0	
No	No	No	Yes	No	Yes	0	0	20	0.03868	0	
No	No	Yes	No	No	Yes	0	0	5	0.00967	0	
No	No	Yes	Yes	No	No	0	0	3	0.00580	0	
Yes	No	No	Yes	No	No	0	0	3	0.00580	0	
Yes	No	No	No	Yes	No	0	0	1	0.00193	0	
No	No	No	Yes	Yes	No	0	0	1	0.00193	0	
No	No	Yes	Yes	No	Yes	0	0	5	0.00967	0	
No	No	No	Yes	Yes	Yes	0	0	270	0.00387	0.13535	
Yes	No	No	Yes	No	Yes	0	0	1	0.00193	0	
Yes	No	Yes	No	No	Yes	0	0	1	0.00193	0	
Yes	No	No	Yes	Yes	No	0	0	1	0.00193	0	
No	No	Yes	Yes	Yes	No	0	0	1	0.00193	0	
Yes	No	Yes	Yes	Yes	No	0	0	1	0.00193	0	
Yes	No	Yes	Yes	No	Yes	0	0	1	0.00193	0	
Yes	No	Yes	No	Yes	Yes	0	0	1	0.00193	0	
Yes	No	No	Yes	Yes	Yes	0	0	1	0.00193	0	
No	No	No	No	No	Yes	1	0.00435	24	0.04642	0.094	
No	No	No	No	No	No	16	0.06957	353	0.68279	0.102	
No	No	No	No	No	Yes	1	0.00435	8	0.01547	0.281	
No	No	Yes	No	No	No	5	0.02174	33	0.06383	0.341	
Yes	No	No	No	No	No	1	0.00435	5	0.00967	0.500	
Yes	No	No	No	Yes	Yes	1	0.00435	1	0.00193	2.25	
No	Yes	Yes	No	No	Yes	3	0.01304	2	0.00387	3.37	
No	Yes	Yes	Yes	No	No	2	0.00870	1	0.00193	4.50	
No	Yes	No	No	No	No	21	0.09130	7	0.01354	6.74	
No	Yes	Yes	No	Yes	Yes	3	0.01304	1	0.00193	6.74	
No	Yes	Yes	Yes	No	Yes	3	0.01304	1	0.00193	6.74	
No	Yes	No	No	No	Yes	11	0.04783	3	0.00580	8.24	
Yes	Yes	No	No	No	No	23	0.10000	4	0.00774	12.92	
No	Yes	Yes	No	No	No	18	0.07826	2	0.00387	20.23	
Yes	Yes	Yes	No	Yes	Yes	11	0.04783	1	0.00193	24.73	
Yes	Yes	No	Yes	Yes	Yes	15	0.06522	1	0.00193	33.72	
No	Yes	No	No	Yes	No	1	0.00435	0	0	∞	
No	Yes	No	Yes	No	No	1	0.00435	0	0	∞	
No	No	Yes	No	Yes	Yes	1	0.00435	0	0	∞	
No	Yes	Yes	No	Yes	No	1	0.00435	0	0	∞	
Yes	Yes	No	No	Yes	No	1	0.00435	0	0	∞	
Yes	Yes	No	Yes	No	No	2	0.00870	0	0	∞	
Yes	Yes	No	No	No	Yes	3	0.01304	0	0	∞	
No	Yes	No	Yes	No	Yes	5	0.02174	0	0	∞	
No	Yes	No	No	Yes	Yes	5	0.02174	0	0	∞	
Yes	Yes	Yes	No	No	No	1495	0.06087	0.41301	0	0	∞
Yes	Yes	Yes	Yes	No	No	1	0.00435	0	0	∞	
Yes	Yes	No	Yes	Yes	No	1	0.00435	0	0	∞	
No	Yes	No	Yes	Yes	Yes	1	0.00435	0	0	∞	
No	No	Yes	Yes	Yes	Yes	1	0.00435	0	0	∞	
Yes	Yes	Yes	No	No	Yes	4	0.01734	0	0	∞	
Yes	Yes	No	Yes	No	Yes	5	0.02174	0	0	∞	
Yes	Yes	No	No	Yes	Yes	8	0.03178	0	0	∞	
Yes	Yes	Yes	Yes	Yes	No	1	0.00135	0	0	∞	
Yes	Yes	Yes	Yes	No	Yes	3	0.01304	0	0	∞	
No	Yes	Yes	Yes	Yes	Yes	3	0.01304	0	0	∞	
Yes	Yes	Yes	Yes	Yes	Yes	33	0.14348	0	0	∞	

FIGURE 9

Likelihood ratios for six diagnostic questions.

From: Collen, Morris, et al., "Automated Multiphasic Screening and Diagnosis" American Journal of Public Health (Volume 54, No. 5, May, 1964), p. 746.

of the asthmatics and 68 percent of the non-asthmatics answered no to all the questions, providing a likelihood of .102 for that set of responses.

These results were used to select a "positive region which locates the combination of questions which screen out as positive the maximum number of patients with asthma with a low error of false negative."³⁰ Recognizing that "for bronchial asthma, an error in diagnosis may not be of immediate serious consequences, a relatively high ... [likelihood ration cutoff] for this test may be selected."³¹

Boolean Algebra

With the development of statistical diagnostic models, it became apparent that diagnosis lent itself to the use of algorithms. It was Ledley and Lusted³² in their article "Reasoning Foundations of Medical Diagnosis" who introduced Boolean Logic to medical diagnosis. They suggested that "the first step in making a logical analysis of this process is to review some symbolism associated with the propositional calculus of symbolic logic. Such symbolism enables the more precise communication of the concepts involved in logical processes."³³ Symbols, x, y, z, \dots , were introduced to represent attributes such as a sign (fever) or disease (pneumonia). Corresponding capital letters were used in statements about these attributes. "For example, y represents the sentence:

The patient has the attribute y . The negation of this statement, the patient does not have the attribute y , is represented by \bar{Y} , where the bar (called negation) over the Y indicates 'not'." ³⁴

Applying Boolean logic, the statement $X \cdot Y$, the intersection of the separate statements X and Y , implies that the patient has both the attribute x and the attribute y . The statement $X + Y$, the union of the statements X and Y , indicates the group which has attribute x , attribute y or both. A third statement, $X \rightarrow Y$, represents the logical statement, "If the patient has attribute x , he then has attribute y ."

The authors generated three functions: ³⁵

1. $E(S(1), \dots, S(n), D(1), \dots, D(m))$ which represents the relationships between diseases the symptoms that comprise medical knowledge,
2. $G(S(1), \dots, S(n))$ which represents the symptoms presented by a patient, and
3. $f(D(1), \dots, D(m))$ which represents the diagnosis as Boolean function of the diseases alone.

"The logical aspect of the medical diagnosis problem is to determine the diseases f such that if medical knowledge E is known, then: if the patient presents symptoms G , he has diseases f . In terms of our symbolic

language, the problem is to determine a Boolean function f that satisfies the following formula:

$$E \rightarrow (G \rightarrow f)$$

This is the fundamental formula of medical diagnosis."³⁶

The difficulty of application of these techniques to medical diagnosis lies not in the theory but in the lack of direct association between cause and effect, or in practical terms, symptom and diagnosis. For example, in a certain disease, there may be a seventy-five percent chance of a symptom being present.

Since "chance" or "probabilities" enter into medical knowledge", then chance or probabilities enter into the diagnosis itself. At present, it may generally be said that specific probabilities are rarely known; medical diagnostic textbooks rarely give numerical values, although they may use words such as "frequently", "very often" and "almost always".³⁷

The first step in discussing a probabilistic analysis of medical diagnosis is to review some definitions and important properties of probabilities. The concept of total probability is concerned with the following question. Suppose we select at random from our population of patients one single patient: what is the chance or total probability, that the patient chosen has certain specified attributes $f(x,y,\dots,z)$? By definition, the total probability is the ratio of the number of patients that have these attributes to the total number of patients from which the random selection is made. If the total number of patients is N , and if $N(f)$ is the number of these patients with attributes f , then the total probability that a patient has attributes f is:

$$P(f) = N(f)/N$$

The conditional probability is analagous to the total probability, where the selection is made only from that subpopulation of patients that have the specified condition.³⁸

The problem now is: Which of these choices is most probable - that is, which of the disease complexes given by the logical diagnosis function f is the patient most likely to have. In terms of conditional probabilities, the probabilistic aspect of the diagnosis problem is to determine the probability that a patient has disease f where it is known that the particular patient presents symptoms G , that is, the probabilistic aspect of medical diagnosis is to evaluate $P(f/G)$ for a particular patient.³⁹

The data upon which the evaluation of $P(f/G)$ is based must, of course, come from medical knowledge. Such medical knowledge is generally also given in the form of conditional probabilities - namely, the probability that a patient having (a) disease complex will have the symptom complex.... It is interesting to note that most diagnostic textbooks discuss the symptoms associated with the disease rather than the reverse, the disease associated with a symptom. The question that naturally arises at this point is: If medical knowledge is in the form $P(G/f)$ that is, probability of having the symptoms given the patient having the disease, then how can we make the diagnosis $P(f/G)$ - that is, the probability of having the disease given the patient having the symptoms.⁴⁰

It is here that Bayes Law, a well-known formula for relating cause and effects in probabilistic terms, becomes important. A derivation of this law appears in the following chapter. Let it suffice at present to state that Bayes Law permits the calculation of the conditional probability of a disease in a given patient utilizing only the likelihood of the symptom complex within each disease.

Bayesian Models

Numerous authors have developed Bayesian models to assist in medical diagnostic problems. Amongst the earliest was one proposed by Warner, et al.⁴¹, in 1961. Choosing congenital heart disease, because "the accuracy of diagnosis ... from clinical symptoms may be checked by cardiac catheterization and/or findings at surgery, and because relatively objective clinical findings may be easily obtained,"⁴² they studied a population ranging in age from one month to over twenty years. Fifty symptoms felt to be of value in the diagnosis of these disorders, and thirty-three diseases were chosen (Figure 10). The symptom-disease matrix (Figure 11) was completed through three sources:

1. a review of previously published data
2. chart reviews of 1035 patients previously seen by this group
3. "estimates based upon the pathologic physiology of the defect in the case of rare defects in which adequate statistics were not available"⁴³

The authors observed that "it is apparent from our experience ... that the most probable diagnosis ... agrees with the actual diagnosis made by physiologic

Table 1.—List of Symptoms to Be Evaluated by Physician

Symptoms	
x ₁	=age 1 mo. to 1 yr.
x ₂	=age 1 to 2 ¹ / ₂ yr.
x ₃	=>20 yrs.
x ₄	=cyanosis, mild
x ₅	=cyanosis, severe (with clubbing)
x ₆	=cyanosis, intermittent
x ₇	=cyanosis, differential
x ₈	=squatting
x ₉	=dyspnea
x ₁₀	=easy fatigue
x ₁₁	=orthopnea
x ₁₂	=chest pain
x ₁₃	=repeated respiratory infections
x ₁₄	=syncope
x ₁₅	=systolic murmur loudest at apex
x ₁₆	=diastolic murmur loudest at apex
x ₁₇	=systolic murmur loudest in left 4th interspace
x ₁₈	=diastolic murmur loudest in left 4th interspace
x ₁₉	=continuous murmur loudest in left 4th interspace
x ₂₀	=systolic murmur with thrill loudest in left 2nd interspace
x ₂₁	=systolic murmur without thrill loudest in left 2nd interspace
x ₂₂	=diastolic murmur loudest in left 2nd interspace
x ₂₃	=continuous murmur loudest in left 2nd interspace
x ₂₄	=systolic murmur loudest in right 2nd interspace
x ₂₅	=diastolic murmur loudest in right 2nd interspace
x ₂₆	=systolic murmur heard best over posterior chest
x ₂₇	=continuous murmur heard best over posterior chest
x ₂₈	=accentuated 2nd heart sound in left 2nd interspace
x ₂₉	=diminished 2nd heart sound in left 2nd interspace
x ₃₀	=right ventricular hyperactivity by palpation
x ₃₁	=forceful apical thrust
x ₃₂	=pulsatile liver
x ₃₃	=absent or diminished femoral pulsation
x ₃₄	=ECG axis more than 110°
x ₃₅	=ECG axis less than 0°
x ₃₆	=R wave greater than 1.2 mv in lead V ₁
x ₃₇	=R' or qR pattern in lead V ₁
x ₃₈	=R wave greater than 2.0 mv in lead V ₆
x ₃₉	=T wave in lead V ₆ inverted (no digitalis)
x ₄₀	=early diastolic murmur loudest at apex
x ₄₁	=late diastolic murmur loudest at apex
x ₄₂	=holo-systolic murmur loudest in left 4th interspace
x ₄₃	=mid-systolic murmur loudest in left 4th interspace
x ₄₄	=holo-diastolic murmur loudest in left 4th interspace
x ₄₅	=early diastolic murmur loudest in left 4th interspace
x ₄₆	=mid-systolic murmur with thrill loudest in 2nd left interspace
x ₄₇	=holo-systolic murmur with thrill loudest in 2nd left interspace
x ₄₈	=mid-systolic murmur without thrill loudest in 2nd left interspace
x ₄₉	=holo-systolic murmur without thrill loudest in 2nd left interspace
x ₅₀	=murmur louder than gr 3/6

Table 2.—List of Diseases Included in Differential Diagnosis

Diseases	
y ₁	=normal
y ₂	=atrial septal defect without pulmonary stenosis or pulmonary hypertension
y ₃	=atrial septal defect with pulmonary stenosis
y ₄	=atrial septal defect with pulmonary hypertension
y ₅	=complete endocardial cushion defect (A-V commun.)
y ₆	=partial anomalous pulmonary venous connections (without atrial septal defect)
y ₇	=total anomalous pulmonary venous connections (supradiaphragmatic)
y ₈	=tricuspid atresia without transposition
y ₉	=Ebstein's anomaly of tricuspid valve
y ₁₀	=ventricular septal defect with valvular pulmonary stenosis
y ₁₁	=ventricular septal defect with infundibular stenosis
y ₁₂	=pulmonary stenosis, valvular (with or without probe-patent foramen ovale)
y ₁₃	=pulmonary stenosis, infundibular (with or without probe-patent foramen ovale)
y ₁₄	=pulmonary atresia
y ₁₅	=pulmonary artery stenosis (peripheral)
y ₁₆	=pulmonary hypertension, isolated
y ₁₇	=aortic-pulmonary window
y ₁₈	=patent ductus arteriosus without pulmonary hypertension
y ₁₉	=pulmonary arteriovenous fistula
y ₂₀	=mitral stenosis
y ₂₁	=primary myocardial disease
y ₂₂	=anomalous origin of left coronary artery
y ₂₃	=aortic valvular stenosis
y ₂₄	=subaortic stenosis
y ₂₅	=coarctation of aorta
y ₂₆	=truncus arteriosus
y ₂₇	=transposed great vessels
y ₂₈	=corrected transposition
y ₂₉	=absent aortic arch
y ₃₀	=ventricular septal defect without pulmonary hypertension
y ₃₁	=ventricular septal defect with pulmonary hypertension
y ₃₂	=patent ductus arteriosus with pulmonary hypertension
y ₃₃	=tricuspid atresia with transposition

FIGURE 10

Symptoms and diseases for a Bayesian congenital heart disease-diagnostic model.

From: Warner, Homer, et al., "A Mathematical Approach to Medical Diagnosis" Journal of the American Medical Association (Volume 177, No. 3, July 22, 1961) p. 179.

Diseases	Incidence	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22
Y1	0.100	01	49	50	01	00	01	00	01	01	10	03	05	05	03	05	01	70	02	07	00	80	01
Y2	.081	10	50	50	02	01	02	00	01	35	50	05	02	40	01	02	02	30	20	02	05	90	02
Y3	.005	30	60	10	20	10	20	00	01	60	70	05	02	10	10	02	02	05	05	02	57	40	01
Y4	.001	10	20	70	30	10	25	00	01	80	90	05	05	15	10	02	02	15	20	02	05	40	20
Y5	.027	20	50	30	15	05	10	00	01	40	50	05	05	30	05	60	15	90	40	02	10	20	10
Y6	.005	10	40	50	01	01	01	00	01	15	20	01	05	05	01	02	02	20	02	02	02	60	05
Y7	.001	20	70	10	65	10	05	00	01	70	80	05	05	20	05	02	02	10	15	10	05	75	05
Y8	.018	50	48	02	30	65	01	00	10	80	90	20	05	15	10	02	05	65	05	05	20	20	02
Y9	.001	10	45	45	22	44	01	00	22	80	80	10	30	15	22	05	25	95	25	05	05	15	02
Y10	.054	40	55	05	25	25	10	00	30	75	90	05	05	10	20	02	02	20	02	05	65	25	02
Y11	.043	10	55	05	30	30	10	00	40	75	90	05	05	10	25	02	02	20	02	05	65	25	02
Y12	.045	20	70	10	01	01	01	00	01	50	65	01	01	01	10	02	02	10	02	05	70	20	02
Y13	.013	20	70	10	01	01	01	00	01	50	65	01	01	01	10	02	02	10	02	02	70	20	02
Y14	.014	30	09	01	10	30	00	00	80	90	99	05	10	05	35	02	02	40	05	05	01	02	02
Y15	.001	05	45	50	01	01	01	00	01	01	01	01	01	01	01	04	01	02	01	01	02	25	02
Y16	.013	10	45	45	01	01	01	00	01	70	95	40	10	10	10	01	01	30	05	01	01	05	30
Y17	.001	30	60	10	05	01	01	00	01	10	10	05	01	10	01	05	10	20	05	60	01	10	05
Y18	.072	20	40	40	01	01	01	00	01	20	20	10	01	10	05	05	15	10	02	50	02	13	05
Y19	.002	20	20	50	45	45	01	00	01	10	20	05	01	01	10	05	02	10	02	20	02	10	02
Y20	.008	20	50	30	01	01	01	00	01	50	50	40	05	10	10	80	20	10	10	02	05	10	02
Y21	.013	70	29	01	01	01	01	00	01	40	50	20	01	05	05	15	02	05	02	02	02	05	02
Y22	.001	70	29	01	01	01	01	00	01	30	30	30	80	15	20	05	01	01	01	01	01	01	01
Y23	.036	10	80	10	01	01	01	00	01	20	30	20	15	01	35	20	02	20	10	02	05	05	01
Y24	.009	10	80	10	01	01	01	00	01	20	30	20	15	01	35	20	02	20	10	02	05	05	01
Y25	.054	10	70	20	01	01	01	00	01	20	30	20	01	01	05	05	01	20	10	02	02	10	01
Y26	.005	50	40	10	30	60	01	00	15	15	30	05	01	20	10	02	02	70	02	02	10	10	02
Y27	.063	80	10	00	20	60	05	10	05	60	70	20	01	05	10	05	02	50	02	02	03	10	02
Y28	.001	30	30	30	30	05	10	00	01	10	20	01	01	01	05	02	70	02	02	05	30	02	02
Y29	.001	60	30	01	01	01	01	80	30	10	50	05	20	01	20	05	02	50	02	02	10	30	02
Y30	.252	15	70	15	01	01	01	00	01	20	30	05	01	15	05	05	20	95	05	02	10	10	05
Y31	.081	30	60	10	30	50	10	00	05	60	70	20	10	20	10	05	01	50	10	02	05	05	25
Y32	.005	30	40	30	01	01	05	50	01	20	30	10	01	10	05	02	02	10	10	02	02	20	10
Y33	.009	10	55	05	50	20	10	00	01	80	90	20	01	30	05	05	10	70	05	02	10	30	10

Symptoms

X23	X24	X25	X26	X27	X28	X29	X30	X31	X32	X33	X34	X35	X36	X37	X38	X39	X40	X41	X42	X43	X44	X45	X46	X47	X48	X49	X50
05	01	00	01	01	15	05	10	03	01	01	01	02	02	02	02	02	01	00	02	70	04	03	00	00	80	05	10
02	01	01	01	01	60	01	80	01	01	01	70	05	05	85	02	02	01	02	01	30	02	20	05	01	90	01	60
03	01	01	01	02	30	15	40	01	05	01	85	05	20	70	02	02	01	01	01	05	01	05	60	01	38	01	70
01	01	01	01	01	95	01	50	01	05	01	85	05	20	70	02	02	01	02	01	15	20	02	05	01	10	01	40
01	01	01	01	01	70	02	40	10	10	01	95	70	05	85	02	02	15	01	85	05	02	20	02	20	20	20	80
05	01	01	10	15	40	02	10	01	01	15	02	02	15	02	02	02	02	02	02	20	02	02	02	02	60	02	30
20	01	01	10	15	85	02	80	01	01	01	90	02	25	75	02	02	02	02	30	10	01	30	05	01	80	02	70
05	01	01	01	01	02	60	01	20	30	01	02	30	02	02	90	10	05	02	50	15	05	02	20	20	20	20	50
05	01	01	01	01	02	35	10	20	10	01	10	02	02	60	02	02	25	25	45	45	25	25	15	15	05	05	50
05	02	02	10	15	10	60	20	01	02	01	95	02	85	10	02	02	02	02	20	05	02	02	60	05	25	05	90
05	02	02	10	15	10	60	20	01	02	01	95	02	85	10	02	02	02	02	20	05	02	02	60	05	25	05	90
10	02	02	01	01	10	60	20	01	05	01	95	02	85	10	02	02	01	01	01	10	02	02	68	01	25	01	80
02	02	02	01	01	10	60	20	01	05	01	95	02	85	10	02	02	01	01	01	10	01	01	68	01	25	01	80
05	02	02	10	10	01	90	20	01	02	01	95	02	85	10	02	02	02	01	30	40	02	05	01	01	02	02	20
01	20	02	50	05	10	02	10	01	01	01	10	02	10	02	02	02	01	01	02	02	01	60	02	01	25	02	60
02	02	02	02	02	95	00	30	01	10	01	95	02	90	05	02	02	01	01	01	30	15	05	02	02	05	02	20
20	02	02	02	02	70	01	20	40	01	01	15	02	02	60	05	10	02	10	20	05	02	02	02	10	05	75	02
85	02	02	03	05	50	01	20	40	02	01	02	10	02	02	50	05	10	02	05	10	02	02	05	02	20	10	85
05	01	01	05	70	05	05	20	01	01	01	05	05	02	02	02	02	02	10	10	02	02	02	02	02	10	10	30
02	02	02	01	01	50	01	20	05	02	01	50	02	10	40	02	02	20	20	10	10	10	10	05	05	10	10	70
02	10	02	01	01	20	02	10	50	02	01	05	10	05	05	40	02	02	10	10	02	02	02	02	02	05	05	10
01	01	01	01	01	26	02	01	05	01	01	05	10	05	05	20	90	01	01	01	01	01	01	01	01	01	01	10
01	95	05	01	01	20	10	01	40	01	05	05	15	02	02	70	15	02	02	02	20	10	02	05	01	05	01	90
01	95	05	01	01	20	10	01	40	01	05	05	15	02	02	70	15	02	02	02	20	10	02	05	01	05	01	90
05	15	10	80	15	19	10	01	30	01	99	05	05	02	02	40	04	01	01	05	20	10	02	02	02	10	05	05
02	02	02	05	10	40	10	30	05	01	01	30	10	40	10	20	05	02	02	40	40	02	02	10	10	10	10	40
02	05	02	01	01	20	10	20	20	02	02	40	20	30	05	20	05	02	02	30	30	02	02	03	03	10	10	50
02	05	02	01	01	20	10	10	10	01	01	20	10	10	10	10	02	02	30	30	02	02	05	05	30	30	60	02
02	05	02	01	01	99	02	40	05	01	10	70	05	80	05	10	05	02	02	30	30	02	02	10	10	30	30	20
01	02	05	01	01	30	02	05	30	01	01	30	10	05	05	15	05	20	02	92	05	05	01	01	10	01	10	85
01	02	05	01	01	90	02	30	05	05																		

studies and observation at surgery at least as often as does the most probable diagnosis estimated by three experienced cardiologists from the same clinical information. Furthermore, the differential diagnosis resulting from solution of the equation is frequently more complete, and, in retrospect, often appears more logical to clinicians than the differential diagnosis listed by each of them prior to seeing the equation's prediction." ⁴⁴

"For most radiologists, having a patient with a primary bone tumor can be considered a rare event, and it is difficult for a radiologist to accumulate in a lifetime a significant experience in the radiological diagnosis of bone tumors."⁴⁵ The work of Lodwick^{46,47} suggests that a Bayesian model can be of great assistance to a radiologist confronted with films suggestive of a bone tumor. To be of value to the clinician, two criteria must be met. The system must be organized so that the clinician is not overwhelmed by the complexity of computer language. Second, definitions must be provided for any symptoms or terms on which a universal opinion does not exist or subjective influence can occur.

Lodwick has presented his Bayesian model in the form of a brochure containing several forms (Figures 12 and 13) which explain the system. Data is entered onto punch cards and a Bayesian analysis performed using Lodwick's

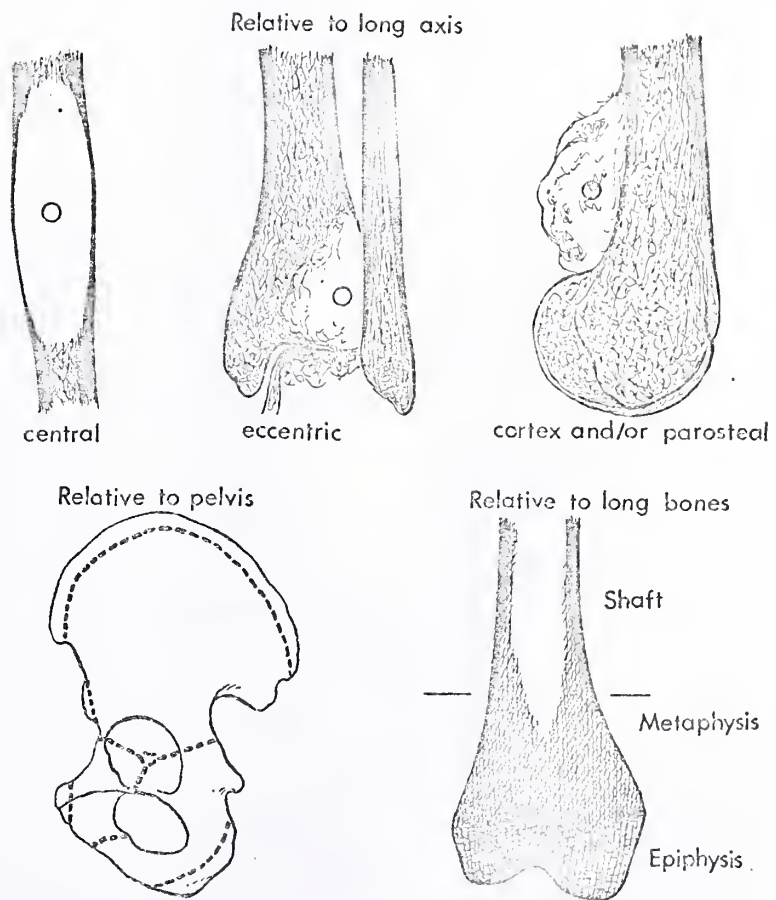


FIGURE 12

FIGURE 12

Guidelines for a Bayesian bone tumor diagnostic model.

From: Lusted, Lee, Introduction to Medical Decision Making
(Springfield, Illinois, Charles Thomas, 1968).

TABLE 1. FORM USED IN COMPUTER-AIDED GRADING AND DIAGNOSIS OF PRIMARY BONE TUMORS
(Delay in diagnosis and 5-year survival are not used by the computer)

Case Number (1-6)							
1. Clinical data		Col.	Yes	No			
Sex	Male	7	1	0			
Age	0-9	8	1	2			
	10-19	9	1	2			
	20-29	10	1	2			
	30-39	11	1	2			
	≥ 40	12	1	2			
Months delay in diagnosis							
	0	13	1	2			
	1-2	14	1	2			
	3-11	15	1	2			
	≥ 12	16	1	2			
Five years survival from diagnosis		17	1	0			
2. Size		mm					
	00	18	1	2			
	01-30	19	1	2			
	31-60	20	1	2			
	61-90	21	1	2			
	91 up	22	1	2			
3. Location							
	Central	23	1	2			
	Eccentric	24	1	2			
	Cortex or surface	25	1	2			
	Pelvis or sacrum	26	1	2			
	Other flat bones, including ribs, skull, mandible	27	1	2			
	Small bones, including patella	28	1	2			
	Tubular bones, including clavicles	29	1	2			
Subdivisions of tubular bones							
Articular surface		30	1	0			
Epiphysis		31	1	0			
Growth plate		32	1	0			
Metaphysis		33	1	0			
Shaft		34	1	0			
4. Tumor matrix mineralization					Col.	Yes	No
Radiolucent		35	1	0			
Cloudy		36	1	0			
Flocculent		37	1	0			
Lumpy		38	1	0			
Solid		39	1	0			
5. Destruction of bone							
Geographic		40	1	0			
Moth-eaten		41	1	0			
Permeated		42	1	0			
Fracture		43	1	0			
Displacement		44	1	0			
Regular margin		45	1	0			
Lobulated margin		46	1	0			
Ragged margin		47	1	0			
Indistinct margin		48	1	0			
Sharp edge		49	1	0			
Smudged edge		50	1	0			
Invasive edge 1 cm		51	1	0			
Inv. edge gr. 1 cm		52	1	0			
Penetration of cortex							
{ None		53	1	2			
{ Partial		54	1	2			
{ Total		55	1	2			
6. Proliferation of bone							
Sclerotic rim		56	1	0			
Mottled		57	1	0			
Endostotic		58	1	0			
Hyperostotic		59	1	0			
Trabeculated		60	1	0			
Buttressed		61	1	0			
{ No expanded cortex		62	1	2			
{ Expanded 10 mm		63	1	2			
{ Expanded 11 mm up		64	1	2			
{ No Codman's triangle		65	1	2			
{ 1 Codman's		66	1	2			
{ 2 Codman's		67	1	2			
{ 3 or more Codman's		68	1	2			
{ No periostosis		69	1	2			
{ Laminated perio.		70	1	2			
{ Amorphous perio.		71	1	2			
{ No spiculation		72	1	2			
{ Regular		72	1	2			
{ Hair-on-end		74	1	2			
{ Velvet		75	1	2			

NOTE: Within each bracketed group only one item may be checked yes.

FIGURE 13

Questionnaire for a Bayesian bone tumor diagnostic model.

From: Lodwick, Gwilym, et al., "Computer Aided Diagnosis of Radiographic Images" Journal of Chronic Diseases (Volume 19, No. 4, April, 1966) p. 489.

"prior distribution" tables. The accuracy of the system is high, with greater than ninety percent agreement with the histological type for nine cell categories. The data analysis was tested using two separate matrices, the first based on data from a Sarcoma Registry, the literature and Lodwick's personal file; the second, employing greater input from Lodwick's personal experience and expectations. Interestingly, the second symptom-disease matrix yielded results five percent better than the first.

A criticism often levelled at computer-dependent diagnostic models is their relative inaccessibility; the results frequently are not available until after the clinical diagnosis has been obtained. The largest Bayesian study to date, a model for the diagnosis of abdominal pain,^{48,49} offers a solution. The study involves an "on-line" facility within the Professorial Surgical Unit at Leeds, England, into which the clinician "logs" the appropriate information. Following the Bayesian computations, the teletypewriter prints a "hardcopy" including a summary of the diagnostic findings, the computer's "diagnoses", a list of additional physical findings to reexamine, and finally, a listing of rarer diseases to be included in the differential diagnosis but too rare to provide a sufficient data base for computer analysis.

The symptoms used in the model (Figure 14) are the standard ones employed in the diagnosis of the "acute abdomen". Seven diagnoses - acute appendicitis, diverticular disease, perforated peptic ulcer, acute cholecystitis, acute intestinal obstruction, acute pancreatitis and non-specific abdominal pain - were chosen for the study since they represented over ninety-five percent of all admissions to this unit with acute abdominal pain. The entity, non-specific abdominal pain, while not an accepted medical entity,⁵⁰ played a significant role in the study. DeDombal⁵¹ has noted that every experienced surgeon encounters a significant number of patients with acute abdominal pain who fit none of the other disease patterns. In some, the pain subsides spontaneously and the patient is sent home undiagnosed, others come to "negative laparotomy", a third group have a medical problem not requiring surgical approach, for example, a urinary tract infection, and still others have psychophysiologic pain. Thus, it is important to have a category for those patients admitted to a surgical unit with acute abdominal pain for whom no surgical diagnosis can be obtained and surgical intervention is inappropriate.

To qualify for the study, the criteria were:

1. patient presented with a chief complaint of abdominal pain

CLINICAL DIAGNOSIS SURVEY

NAME [REDACTED]		SERIAL NO
REGISTRATION NO. [REDACTED]	CODE 3703	MRC PRO
SEX FEMALE		AA
AGE 46		



<p>A. PRESENTING SYMPTOM PAIN</p> <p>B. PAIN</p> <p>1. Site at onset </p> <p>2. Site at present </p> <p>3. Severity SEVERE</p> <p>4. Aggrav. fact. NIL</p> <p>5. Relief fact. NIL</p> <p>6. Progress SAME</p> <p>7. Duration 16 HRS</p> <p>8. Type - at onset INTERMITTENT</p> <p>9. Type - at present INTERMITTENT</p> <p>C. OTHER SYMPTOMS</p> <p>1. Nausea YES</p> <p>2. Vomiting YES</p> <p>3. Appetite DECREASED</p> <p>4. Pre. indigestion YES</p> <p>5. Jaundice YES</p> <p>6. Bowels NORMAL</p> <p>7. Micturition NORMAL</p> <p>8. Previous Pain YES</p>	<p>E. GENERAL EXAMINATION</p> <p>1. Mood. DISTRESSED</p> <p>2. Colour. FLUSHED</p> <p>3. Temp.</p> <p>4. Pulse</p> <p>5. BP</p> <p>F. ABDOMINAL INSPECTION</p> <p>1. Movement NORMAL</p> <p>2. Scars NO</p> <p>3. Distension NO</p> <p>G. ABDOMINAL PALPATION</p> <p>1. Tenderness GENERAL</p> <p>2. Rebound NO</p> <p>3. Guarding NO</p> <p>4. Rigidity YES</p> <p>5. Swelling NIL</p> <p>6. Murphy NEGATIVE</p> <p>H. ABDOMINAL AUSCULTATION ABSENT</p> <p>I. RECTAL EXAMINATION TENDER LEFT SIDE</p> <p>J. CLINICIANS PRE OP D 1, 2, 3</p>
---	---

FIG. 3—Example of case history, showing form on to which patient data are copied for later entry into the computer.

FIGURE 14

Questionnaire for a Bayesian abdominal pain diagnostic model.
 From: Horrocks, J.C., et al., "Computer-Aided Diagnosis"
British Medical Journal (Volume 2, April 1, 1972) p. 6.

2. the pain was of less than one week's duration
3. the admission was an emergency procedure through the emergency room
4. the patient had not previously entered into the study
5. the patient was capable of giving a history (this excluded only two patients, a baby aged two weeks and a comatose patient)
6. a diagnosis was eventually made (excluding one patient for whom a histological diagnosis was indecisive).

The results of the study are striking. Compared to each clinician, the computer yielded consistently better results, as shown in Figure 15. The diagnosis of acute appendicitis, the most frequent cause of the acute abdomen, provided data of great interest. Of the eighty-five patients with acute appendicitis, the computer diagnosed 84. The Senior clinicians recognized only 75 and classified six into the non-operative category, often delaying surgery several hours. While the computer erred in diagnosing nine additional patients as having acute appendicitis, the clinicians (without knowledge of the computer's diagnosis) performed more than twenty negative

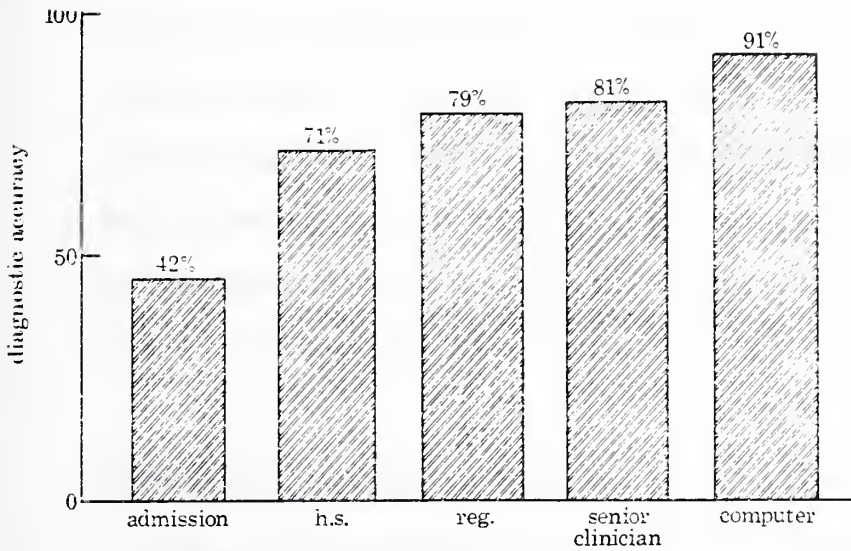


FIGURE 2. Comparison of overall human and computer-aided diagnostic accuracy.

FIGURE 15

From: deDombal, F.T., "Surgical Diagnosis assisted by a Computer" Proceedings of the Royal Society of London (Volume 184, 1973) p. 437.

laparotomies, including those nine. DeDombal states, "We cannot resist pointing out two facts. First, the cost of performing a 'negative laparotomy' and keeping the patient in the hospital for a week postoperatively is rarely less than £200 (\$500); and the system would have obviated the need for 34 such operations [in a larger series]. More important perhaps, diagnostic delay which results in perforation of an appendix increases the mortality of appendectomy tenfold; the system's predictions would have helped to minimize such delay."⁵²

Other diagnostic models

Two other diagnostic models are worthy of comment although it is difficult to classify them into the previous categories. In a monograph entitled "Diagnostic Computers", Caceres and Rikli⁵³ describe an operational computer program to "read" electrocardiograms and propose diagnoses. The protocol involves:

1. minimization of "noise" from the circuitry
2. choosing a fiducial point
3. determination of significant reference points
4. pattern recognition routines.

Once the wave patterns are recognized, the data is "condensed to those values significant to diagnosis by

a discriminant function analysis,"⁵⁴ and tested through multiple combinations. For example, "the diagnosis of 'acute diaphragmatic myocardial infarction' is pursued by checking for the presence of a pre-determined positive value for Q duration, for negative T waves, and for elevated ST segments".⁵⁵ To obtain the final EKG diagnosis, "All of the resultant diagnostic statements are initially considered as only tentative. Those which are redundant or secondary to diagnosis of greater strength are discarded".⁵⁶

To compare the program to a clinician, 750 routine EKGs, of which approximately twenty-five percent were normal, were presented to an independent cardiologist (one not aware of the criteria employed by the program) and to the computer. There was complete agreement in about 73% of the readings. In 26%, clinician and computer agreed that the tracing was abnormal, but disagreed on the nature of the abnormality. In 0.8 percent of the cardiograms, the computer read as normal something considered abnormal by the clinician. The vast majority of these were felt to be "analog magnetic tape playback, digitilization and artifacts".⁵⁷

The authors foresaw improvement in diagnostic accuracy with technical refinements and additional arrhythmia routines. Considering the estimated cost of between two and four dollars per reading, with improvements the program

could be viable for clinical use.

Diagnostic Criteria

The diagnosis of certain medical diseases, particularly collagen disorders, is difficult because of the variety of presentations of the same disease. To surmount this problem, it has been necessary to establish diagnostic criteria. At the meetings of the American Rheumatism Association in 1954, there was a demand for precise diagnostic criteria for rheumatoid arthritis. In 1956, a committee headed by Ropes⁵⁸ proposed criteria for the classes of "definite, probable and possible rheumatoid arthritis." After two years of use, the lack of rigor in the criteria was criticized and revised criteria⁵⁹ were published, adding a fourth category, classical rheumatoid arthritis.

The criteria (Figure 16), while controversial, were felt to have a number of benefits. Most importantly, medical school faculty "have found these criteria helpful as a framework from which to discuss the disease. This has proved true even for some ... who disagree with the criteria or some portion thereof, for they still have a framework from which to operate and a specific opportunity can then be made to express differing points of view."⁶⁰

This diagnosis requires two of the following criteria, and the total duration of joint symptoms must be at least three weeks. (Any one of the features listed under EXCLUSIONS will exclude a patient from this category.)

1. Morning stiffness.
2. Tenderness or pain on motion (observed by a physician) with history of recurrence or persistence for three weeks.
3. History or observation of joint swelling.
4. Subcutaneous nodules (observed by a physician).
5. Elevated sedimentation rate or C-reactive protein.
6. Iritis

EXCLUSIONS

1. The typical rash of *disseminated lupus erythematosus* (with butterfly distribution, follicle plugging, and areas of atrophy).
2. High concentration of *lupus erythematosus* cells (4 or more in 2 smears prepared from heparinized blood incubated not over 2 hrs.).
3. Histologic evidence of *periarteritis nodosa* with segmental necrosis of arteries associated with nodular leukocytic infiltration extending perivascularly and tending to include many eosinophils.
4. Weakness of neck, trunk, and pharyngeal muscles or persistent muscle swelling of *dermatomyositis*.
5. Definite *scleroderma* (not limited to the fingers).
6. A clinical picture characteristic of *rheumatic fever* with migratory joint involvement and evidence of endocarditis, especially if accompanied by subcutaneous nodules, erythema marginatum, or chorea. (An elevated antistreptolysin titer will not rule out the diagnosis of rheumatoid arthritis.)
7. A clinical picture characteristic of *gonorrhea* with acute attacks of swelling, redness, and pain in one or more joints, especially if relieved by colchicine.
8. *Typhi*.
9. A clinical picture characteristic of acute *infectious arthritis* of bacterial or viral origin with an acute focus of infection, or in close association with a disease of known infectious origin; chills; fever; and an acute joint involvement, usually migratory initially (especially if there are organisms in the joint fluid or if there is response to antibiotic therapy).
10. *Tubercle bacilli* in joints or histologic evidences of joint tuberculosis.
11. A clinical picture characteristic of *Reiter's syndrome* with urethritis and conjunctivitis associated with acute joint involvement, usually migratory initially.
12. A clinical picture characteristic of the *shoulder-hand syndrome*: unilateral involvement of shoulder and hand with diffuse swelling of the hand followed by atrophy and contractures.
13. A clinical picture characteristic of *hypertrophic pulmonary osteoarthropathy* with clubbing of fingers and/or hypertrophic periostitis along the shafts of the long bones, especially if an intrapulmonary lesion is present.
14. A clinical picture characteristic of *neuroarthropathy* with condensation and destruction of bones of involved joints and with associated neurologic findings.
15. *Hyperuricemic acid* in the urine, detectable grossly with alkalization.
16. Histologic evidence of *carcinoid* or positive Kveim test.
17. *Multiple myeloma* as evidenced by marked increase in plasma cells in the bone marrow, or hence Jones protein in the urine.
18. Characteristic skin lesions of *erythema nodosum*.
19. *Leukemia* or *lymphoma* with characteristic cells in peripheral blood, bone marrow, or tissues.
20. *Syganactobulbaria*.

This diagnosis requires seven of the following criteria. In criteria 1 through 5, the joint signs or symptoms must be continuous for at least six weeks. (Any one of the features listed under EXCLUSIONS will exclude a patient from this category.)

1. Morning stiffness.
2. Pain on motion, or tenderness in at least one joint (observed by a physician).
3. Swelling (soft tissue thickening or fluid—not bony overgrowth alone) in at least one joint (observed by a physician).
4. Swelling (observed by a physician) of at least one other joint (any interval free of joint symptoms between the two joint involvements may not be more than 3 months).
5. Symmetric joint swelling (observed by a physician) with simultaneous involvement of the same joint on both sides of the body (bilateral involvement of metacarpophalangeal, metacarpophalangeal, or metatarsophalangeal joints is acceptable with absolute symmetry). Terminal phalangeal joint involvement will not satisfy this criterion.
6. Subcutaneous nodules (observed by a physician) over bony prominences, on extensor surfaces, or in intra-articular regions.
7. X-ray changes typical of rheumatoid arthritis (which must include at least bony decalcification localized to or greatest about the involved joints and not just degenerative changes). Degenerative changes do not exclude patients from any group classified as rheumatoid arthritis.
8. Positive agglutination test—demonstration of the "rheumatoid factor" by any method that, in two laboratories, has been positive in not over 5 percent of normal controls—or positive streptococcal agglutination test.
9. Poor mucin precipitate from synovial fluid (with shreds and cloudy solution).
10. Characteristic histologic changes in synovial membrane with three or more of the following: marked villous hypertrophy; proliferation of superficial synovial cells, often with palisading; marked infiltration of chronic inflammatory cells (lymphocytes or plasma cells predominating) with tendency to form "lymphoid nodules;" deposition of compact fibrin, either on surface or interstitially; foci of cell necrosis.
11. Characteristic histologic changes in nodules showing granulomatous foci with central zones of cell necrosis, surrounded by proliferated fixed cells, and peripheral fibrosis and chronic inflammatory cell infiltration, predominantly perivascular.

Definite Rheumatoid Arthritis

This diagnosis requires five of the foregoing criteria. In criteria 1 through 5, the joint signs or symptoms must be continuous for at least six weeks. (Any one of the features listed under EXCLUSIONS will exclude a patient from this category.)

Probable Rheumatoid Arthritis

This diagnosis requires three of the above criteria. In at least one of the criteria, 1 through 5, the joint signs or symptoms must be continuous for at least six weeks. (Any one of the features listed under EXCLUSIONS will exclude a patient from this category.)

FIGURE 16

Diagnostic criteria for rheumatoid arthritis.

From: Lusted, Lee, Introduction to Medical Decision Making (Springfield, Illinois, Charles Thomas, 1968) p. 65-67.

FOOTNOTES

1. F. A. Nash, "Differential Diagnosis" The Lancet (April 17, 1954, Volume 266, No. 6816) pp. 874-875.
2. Martin Lipkin, James Hardy, "Differential Diagnosis of Hematological Diseases Aided by Mechanical Correlation of Data" Science (March 22, 1957, Vol. 125) pp. 551-552.
3. Ibid., p. 551.
4. Martin Lipkin, James Hardy, "Mechanical Correlation of Data in Differential Diagnosis of Hematological Diseases" Journal of the American Medical Association (January 11, 1958, Vol. 166, No. 2) pp. 113-125.
5. R. S. Casey, J. W. Perry, Punched Cards, Their Application to Science and Industry (New York, Reinhold, 1951) p. 11.
6. Lipkin, Hardy, "Mechanical", Op. Cit., pp. 116-117.
7. Lewis Barness, Walter Tunnessen, et al., "Computer Assisted Diagnosis in Pediatrics" American Journal of Diseases of Children (June, 1974, Volume 127) pp. 852-858.
8. Ibid., p. 853.
9. Ibid., p. 858.
10. Jerome Krovetz, Ira Gessner, Gerold Schiebler, Handbook of Pediatric Cardiology (New York, Hoebner Medical Division, Harper and Row, 1969) p. 367.
11. Benjamin Kleinmuntz, "Diagnostic Problem Solving by Computer" Japanese Psychological Research (1965, Vol. 7, No. 4) pp. 189-194.
12. Ibid., p. 190.
13. Loc. Cit.
14. Loc. Cit.
15. Ibid., pp. 190, 191.
16. Don Brand, Henry Dove, Alan Meyers, "A Technique for Analyzing Clinical Data to Provide Patient Management Guideline: A Study of Meningitis in Children" Submitted for publication.
17. Martin Lipkin, "The Likelihood Concept of Differential Diagnosis" Perspectives in Medicine and Biology (Summer, 1964) pp. 485-497.
18. Ibid., p. 486.
19. Ibid., p. 487.
20. Ibid., p. 488.

21. J. Crooks, I. Murray, E. Wayne, "Statistical Methods Applied to the Clinical Diagnosis of Thyrotoxicosis" Quarterly Journal of Medicine (April, 1959, Vol.28, No. 110) pp. 211-234.
22. Ibid., p. 212.
23. Loc. Cit.
24. Ibid., p. 214.
25. Morris Collen, Leonard Rubin, Jerzy Neyman, et al., "Automated Multiphasic Screening and Diagnosis" American Journal of Public Health (May, 1964, Vol. 54, No. 5) pp. 741-750.
26. Ibid., p. 741.
27. Jerzy Neyman, "Outline of Statistical Treatment of the Problem of Diagnosis" Public Health Reports (October 3, 1947, Vol. 62) pp. 1449-1456.
28. Collen, Op. Cit. p. 743.
29. Ibid., pp. 745-747.
30. Ibid., p. 747.
31. Ibid., p. 747.
32. Robert Ledley, Lee Lusted, "Reasoning Foundations of Medical Diagnosis" Science (July 3, 1959, Volume 130, No. 3366) pp. 9-21.
33. Ibid., p. 10.
34. Loc. Cit.
35. Ibid., p. 11.
36. Loc. Cit.
37. Ibid., p. 13.
38. Loc. Cit.
39. Ibid., p. 14.
40. Loc. Cit.
41. Homer Warner, Alan Toronto, George Veasy, Robert Stephenson, "A Mathematical Approach to Medical Diagnosis" Journal of the American Medical Association (July 22, 1961, Volume 177, No. 3) pp. 171-183.
42. Ibid., p. 179.
43. Ibid., p. 180.
44. Ibid., p. 183.
45. Lee Lusted, Introduction to Medical Decision Making (Charles Thomas, Springfield, Illinois, 1968).

46. Gwilym Lodwick, Cosmo Haun, et al., "Computer Diagnosis of Primary Bone Tumors" Radiology (February, 1963, Vol. 80) pp. 273-275.
47. Gwilym Lodwick, Arch Turner, et al., "Computer-Aided Analysis of Radiographic Images" Journal of Chronic Diseases (April, 1966, Vol. 19, No. 4) pp. 485-496.
48. Jane Horrocks, A. P. McCann "Computer-Aided Diagnosis" Description of an Adaptable System and Operational Experience with 2034 Cases" British Medical Journal (April 1, 1972, Vol. 2) pp. 5-9.
49. F. T. deDombal, D. J. Leaper, "Computer-Aided Diagnosis of Acute Abdominal Pain" British Medical Journal (April 1, 1972, Vol. 2) pp. 9-13.
50. J. Shepherd, "Correspondence" British Medical Journal (Vol. 2, 1972) p. 347.
51. F. T. deDombal, "Surgical Diagnosis Assisted by a Computer" Proceedings of the Royal Society of London (1973, Vol. 184) pp. 433-440.
52. Ibid., p. 439.
53. Cesar Caceres, Arthur Rikli, Diagnostic Computers (Springfield, Illinois, Charles Thomas, 1969).
54. Ibid., p. 39.
55. Ibid., pp. 41-42.
56. Ibid., p. 42.
57. Ibid., p. 44.
58. Marian Ropes, Granville Bennett, Sidney Cobb, et al., "Proposed Diagnostic Criteria for Rheumatoid Arthritis" Annals of Rheumatic Disease (March 25, 1957, Vol. 16) pp. 118-125.
59. Committee of the American Rheumatism Association. "1958 Revision of Diagnostic Criteria for Rheumatoid Arthritis" Arthritis and Rheumatism (1959, Vol. 2) pp. 16-20.
60. Sidney Cobb, "On the Development of Diagnostic Criteria" Arthritis and Rheumatism (Vol. 13, February, 1960) pp. 93.

CHAPTER III

Bayes Theorem

It was in 1763 that Bayes¹ first proposed his theorem of prior distributions. Despite numerous onslaughts from the world of mathematics, it has appeared repeatedly in scientific literature and has recently entered the medical literature, dealing with problems of diagnosis or prognosis.

Prior to defining the theorem, it is necessary to define a number of mathematical terms. These appear in Figure 17.

In mathematical theory, the probability of an event is a number between zero (an impossible event) and one (a certainty). Typically, the result lies between and can be defined as:

$$P(D) = N(D)/N,$$

the ratio of the number of people with the disease to the entire population. Similarly, the probability of a symptom can be defined as:

$$P(S) = N(S)/N.$$

FIGURE 17

N	Number of people in the population
$N(D)$	Number of people in this population with disease D
$N(S)$	Number of people with symptom S
$N(D \cap S)$ or $N(S \cap D)$	Number of people with both the disease D and the symptom S
$P(D)$	Probability of disease D
$P(S)$	Probability of symptom S
$P(S \cap D)$ or $P(D \cap S)$	Probability of both the disease D and the symptom S
$P(S_1, S_2, \dots, S_n)$	Probability of a set of symptoms S_1, S_2, \dots, S_n
$P(D/S)$	Probability of disease D, given symptom S
$P(S/D)$	Probability of symptom S, given disease D
$P(D/S_1, S_2, \dots, S_n)$	Probability of disease D, given a set of symptoms, S_1, S_2, \dots, S_n

The probability of a symptom being present, given a disease, is proportional to the number of patients having both the symptom and the disease, and inversely to the number of patients having the disease, that is:

$$P(S/D) = N(S \cap D)/N(D).$$

In the special case $N(D) = 0$, the formula is invalid and $P(S/D)$ is undefined. Dividing both numerator and denominator by N yields:

$$P(S/D) = \frac{N(S \cap D)/N}{N(D)/N}$$

This can be simplified to:

$$P(S/D) = P(S \cap D)/P(D) \quad (1)$$

A similar derivation yields the result:

$$P(D/S) = P(S \cap D)/P(S)$$

Crossmultiplying yields:

$$P(S \cap D) = P(D) \cdot P(S/D)$$

$$P(S \cap D) = P(S) \cdot (P(D/S))$$

Thus:

$$P(D) \times P(S/D) = P(S) \times P(D/S)$$

In a different form (excluding the case in which $P(S) = 0$:

$$P(D/S) = \frac{P(D) \cdot P(S/D)}{P(S)}$$

This is Bayes Law. It states that the probability of a disease being present is proportional to three factors: directly to the incidence of the disease as well as the probability of the symptoms within the disease, and inversely to the probability of the symptoms with the population. It is important to note that no assumptions have been made regarding the nature of the distributions and the inter- or independence of symptoms or diseases. Other than the assumption that the denominator cannot be zero, it has been demonstrated that the probability of a disease being present given a set of symptoms can be universally defined by three pieces of data - the probability of a symptom within a disease and the probability of each the symptom and disease within the population being studied (a community, country or clinic population, for example). An example offered by Mount and Evans² can illustrate the use of Bayes Law. A population of one thousand persons is distributed with regard to a symptom and disease as follows:

1. 100 patients have both the disease and symptom
2. 200 patients have the disease but not the symptom
3. 300 patients have neither the disease nor the symptom

4. 400 patients do not have the disease but have the symptom.

These statements are equivalent to:

1. $N(S \cap D) = 100$
2. $N(\bar{S} \cap D) = 200$
3. $N(\bar{S} \cap \bar{D}) = 300$
4. $N(S \cap \bar{D}) = 400$

in which \bar{S} and \bar{D} imply respectively, those without the symptom and those without the disease. It is simple to compute:

$$\begin{aligned} P(D) &= P(D \cap S) + P(D \cap \bar{S}) \\ &= 100/1000 + 200/1000 \\ &= 0.3 \end{aligned}$$

and,

$$\begin{aligned} P(S) &= 100/1000 + 400/1000 \\ &= 0.5 \end{aligned}$$

Using Equation (1):

$$\begin{aligned} P(S/D) &= P(S \cap D)/P(D) \\ &= 0.1/0.3 \\ &= 0.33 \end{aligned}$$

To determine the probability of a patient having this disease, with the knowledge that the patient has the symptom, one can use Bayes Law:

$$\begin{aligned}
 P(D/S) &= P(D)P(S/D)/P(S) \\
 &= 0.3 \times 0.33/0.9 \\
 &= 0.2.
 \end{aligned}$$

This value could be derived by direct examination of the data with complete agreement.

Diseases are not typically represented by one symptom but more likely by a set of symptoms, physical findings or laboratory abnormalities. Symptoms are better seen as complexes. Instead of $P(S)$, it is more appropriate to employ the term $P(S_1, S_2, \dots, S_n)$ and Bayes Law becomes:

$$P(D/S_1, S_2, \dots, S_n) = \frac{P(D) \cdot P(S_1, S_2, \dots, S_n/D)}{P(S_1, S_2, \dots, S_n)}$$

In addition, the absence of a symptom is as important to consider as its presence. The probability of \bar{S}_j - the symptom S_j being absent - is equal to $1 - P(S_j)$.

In the simplest of clinical situations, one is concerned with the presence or absence of a single disease. This is equivalent to $P(D) + P(\bar{D}) = 1$, that is, the patient either has, or does not have the disease; no other possibilities are tenable. If two or more diseases are considered, the patient must always have some disease or be normal:

$$P(D_1) + P(D_2) + \dots + P(D_n) + P(\text{normal}) = 1.$$

For the sake of simplification, the state normal is either included as an additional disease state or ignored.

The terms $P(S)$ can also be expanded, to:

$$P(S) = P(S/D_1)P(D_1) + P(S/D_2)P(D_2) + \dots + P(S/D_m)P(D_m)$$

Combining these, Bayes Law becomes:

$$P(D_i/S_1, \bar{S}_2, \dots, S_n) = \frac{P(S_1, \bar{S}_2, \dots, S_n/D_i)P(D_i)}{P(S_1, \bar{S}_2, \dots, S_n/D_1)P(D_1) + P(S_1, \bar{S}_2, \dots, S_n/D_2)P(D_2) + \dots + P(S_1, \bar{S}_2, \dots, S_n/D_i)P(D_i) + \dots + P(S_1, \bar{S}_2, \dots, S_n/D_m)P(D_m)}$$

still without having made any assumptions.

To apply Bayes Law to clinical situations, it is necessary to know only the following information:

1. the incidence (or prevalence, if appropriate) of the disease within the population being studied, and,
2. the probability of the symptom complex being present within each of the diseases under consideration.

The first is accessible through the medical literature or public health data. If the population under consideration is a clinic group, the data can be predicted from personal experience. The second is somewhat more complex. Medical literature generally contains information in the form of $P(S/D)$ for each individual symptom. Only rarely does the literature propose to give $P(S_1, S_2/D)$, the occurrence

of two symptom complexes within a disease. Even rarer does it propose to deal with more extensive symptom complexes for any disease.

To obtain sufficient data would require a review of a large number of cases, proportional to $m \times 2^n$ for m disease and n symptoms, each having only two possible states, present or absent. As the number of symptoms becomes large or the states for any one symptom increases (for example, a symptom graded as absent, mild, moderate or severe) the task becomes untenable: the required sample would quickly exceed all human population, past or present.

To simplify this task, this study, as well as all the Bayesian studies in the literature, assumes the independent sorting of symptoms within each disease. If independent, the incidence of a set of symptoms can be expanded to:

$$P(S_1, S_2, \dots, S_n) = P(S_1)P(S_2) \dots P(S_n)$$

and Bayes Law becomes:

$$P(D_i/S_1, S_2, \dots, S_n) = \frac{P(D_i)P(S_1/D_i)P(S_2/D_i) \dots P(S_n/D_i)}{P(D_1)P(S_1/D_1)P(S_2/D_1) \dots P(S_n/D_1) + P(D_2)P(S_1/D_2)P(S_2/D_2) \dots P(S_n/D_2) + \dots + P(D_m)P(S_1/D_m)P(S_2/D_m) \dots P(S_n/D_m)}$$

Theoretically, this cannot be true, as demonstrated by Lusted. "Two symptoms S_1 and S_2 relevant to a disease D

cannot be independent of each other unconditionally since the truth or falsehood about the hypothesis of what disease causes S_1 and S_2 will cause the symptoms, if they are indeed diagnostically relevant, to co-vary."³ Warner⁴ dealt with this problem through the mathematical demonstration that for a symptom to be truly independent of another would require a uniform distribution through the population. Consequently, the symptom would be of no diagnostic value. The use of non-independent symptoms is valid if the correlation "is due only to the non-uniform distribution of x_b in diseases y_1, y_2, \dots, y_k and not due to a direct causal relationship between x_a and x_b ."⁵

Lodwick states "Our version of Bayes rule assumes that our predictor findings are independent, which is to say that information regarding the presence or absence of one finding does not provide information regarding the probability of the presence or absence of any other findings. One can at best only demonstrate correlation between findings in a given sample. The absence of correlation merely suggests the possibility of independence; it can by no means be accepted as proof. Thus, we would appear to be on theoretically shaky grounds to use Bayes theorem in the form we use it."⁶

In practice, independence of symptoms can be tested through chi-squared analysis. If present, highly correlated symptoms can be combined; symptoms, which are causally related, can be eliminated. For example, it would be in error in a congenital heart disease model to employ both cyanosis and pO_2 levels; their correlation, while dependent on other factors, is too high to allow their simultaneous use.

Alternatively, a complicated mathematical procedure can be employed. Brunk and Lehr,⁷ using the Gram Schmidt Orthogonalization Procedure, discard symptoms with a high correlation. Those with weaker correlations selected individually and a linear function of one variable is subtracted from another so that they are no longer correlated. Then another variable is selected, the same function performed, until a set of uncorrelated variables is obtained.

Despite this limitation, Lusted warns against ignoring Bayes Law. "How seriously should you take your inability to prove that symptoms are independent? One answer is that you should not take it seriously enough to discourage you from using Bayesian procedures to study medical diagnosis. You can assume that symptoms are conditionally independent even when you have reason to suppose they are not. . . ., but you should not take the situation too glibly. Watch for violations of conditional independence

which are so severe that they may lead you to major errors in diagnosis"⁸

To illustrate the use of Bayes Law, a symptom disease matrix is shown below.

	Incidence	$P(S_1)$	$P(S_2)$	$P(S_3)$
D	0.25	.2	.3	.9
\bar{D}	0.75	.7	.4	.4

A patient under consideration has symptoms S_1 and S_3 but not S_2 . Using Bayes Law:

$$\begin{aligned}
 P(D/D_1, \bar{S}_2, S_3) &= \frac{P(D)P(S_1, \bar{S}_2, S_3/D)}{P(D)P(S_1, \bar{S}_2, S_3/D) + P(\bar{D})P(S_1, \bar{S}_2, S_3/\bar{D})} \\
 &= \frac{P(D)P(S_1/D)P(\bar{S}_2/D)P(S_3/D)}{P(D)P(S_1/D)P(\bar{S}_2/D)P(S_3/D) + P(\bar{D})P(S_1/\bar{D})P(\bar{S}_2/\bar{D})P(S_3/\bar{D})}
 \end{aligned}$$

Remembering that $P(\bar{S})$ is equal to $1 - P(S)$:

$$\begin{aligned}
 P(D/S_1, \bar{S}_2, S_3) &= \frac{(.25)(.2)(1-.3)(.9)}{(.25)(.2)(1-.3)(.9) + (.75)(.7)(1-.4)(.4)} \\
 &= 0.2
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 P(\bar{D}/S_1, \bar{S}_2, S_3) &= \frac{P(\bar{D})P(S_1, \bar{S}_2, S_3/\bar{D})}{P(\bar{D})P(S_1, \bar{S}_2, S_3/\bar{D}) + P(D)P(S_1, \bar{S}_2, S_3/D)} \\
 &= \frac{(.75)(.7)(1-.4)(.4)}{(.75)(.7)(1-.4)(.4) + (.25)(.2)(1-.3)(.9)} \\
 &= 0.8
 \end{aligned}$$

Thus, the probability of the disease, given these symptoms, is 0.2 (twenty percent) and the probability of the disease not being present is 0.8 (eighty percent).

One final comment on Bayes Law relates to the simultaneous occurrence of two diseases. The data in the symptom-disease matrix relates only to each individual disease. Bayes Law does not permit assumptions regarding the simultaneous occurrence of two diseases: this case must be handled as a new disease pattern. Neither a ventricular septal defect nor pulmonary stenosis alone can cause cyanosis, yet in combination they are likely to do so. A Bayesian model employing the two diseases separately is incapable of diagnosing the combined condition: to diagnosis a VSD with pulmonary stenosis would require data on this third disease.

In the following chapters, a model employing Bayes Law for the diagnosis of neonatal heart disease will be described. In considering the model and its use of Bayes Law, it is important to remember that independence of variables was assumed to allow one to expand Bayes Law to its final form. Beyond this, Bayes Law is of universal validity: it allows its user to predict the likelihood of a disease from a set of symptoms. The only information required for this prediction is the incidence of each disease under consideration and the probability of each symptom within each disease.

FOOTNOTES

1. T. Bayes, "Essay towards solving a problem in the doctrine of chances" Philosophical Transactions of the Royal Society (Volume 53, 1763) pp. 370-418 (Reprinted in Biometrika (Volume 45, 1958) pp. 293-315).
2. James Mount, John Evans, "Computer-Aided Diagnosis-A Simulation Study" Proceedings of the Fifth IBM Medical Symposium (1963) pp. 113-127.
3. Lee Lusted, Introduction to Medical Decision Making (Charles Thomas, Springfield, Illinois, 1968).
4. Homer Warner, Alan Toronto, George Veasy, Robert Stephenson, "A Mathematical Approach to Medical Diagnosis" Journal of the American Medical Association (July 22, 1961, Volume 177, No. 3) p. 178.
5. Ibid., p. 179.
6. Gwilym Lodwick, "Computer-Aided Diagnosis in Radiology. A Research Plan" Investigative Radiology (January, February, 1966, Vol. 1) pp. 75-76.
7. H. D. Brunk, J. L. Lehr, "An Improved Bayes Method for Computer Diagnosis" Proceedings of Conference on the Use of Computers in Radiology (University of Chicago, Center for Continuing Education, October 22, 1966).
8. Lusted, Op. Cit., p. 18.

CHAPTER IV

METHODOLOGY

With the development of high-speed electronic computers, it became practical to apply Bayes Law to the solution of complex clinical problems. As mentioned previously, models have been developed for the diagnosis of congenital heart disease, the acute abdomen, and radiological evaluation of bone tumors. In addition, studies have been performed in the areas of thyroid disease¹ and the diagnosis of Cushing's syndrome.²

Although the area of congenital heart disease was well studied by Warner, Veasy and Toronto, their diagnostic model involved patients ranging in age from one month to twenty years. Another population, those who display their initial symptoms in the period between birth and one month were excluded. A great number of patients with congenital heart disease, particularly those with the most severe disorders, are seen in the neonatal period. More than one-fourth of all infants with congenital heart defects die before the age of one month, most within the first week.³

It has been the purpose of this study to apply the techniques of Bayesian analysis to cardiological diagnosis in the newborn period. The area is well suited to a

predictive model: diagnosis is based primarily on physical findings and radiographic and electrocardiographic data rather than the more subjective history. Also, as the majority of patients, other than those with small PDAs and pulmonary disease, are catheterized, clinical diagnoses can be checked by objective data.

Development of a Bayesian statistical model requires a number of steps. One must specify the clinical problem and identify the patient population. Next, one must decide upon the relevant symptoms, physical findings and laboratory, electrocardiographic and radiographic data and prepare a comprehensive list of diseases which occur in the population. The Bayesian system then depends upon completion of the symptom-disease matrix: data about the incidence of each disease within the population, and the occurrence of each symptom within each disease. Using a computer program to tabulate the large quantity of data under consideration, a number of cases must be analyzed to determine the validity of the system.

The clinical problem in this thesis is congenital heart disease in the neonate. The population which has been selected was arbitrarily limited to any patient admitted to the Yale-New Haven Hospital Newborn Special Care Unit. While this may exclude a few neonates admitted to a different hospital unit or in rare cases may include

patients beyond the first month of life, it has been necessary to adhere to these limits since the symptom-disease was constructed with regard to that population. Each of the other components will be discussed in its approach section.

Symptoms

To assess the symptoms relevant to the diagnosis of congenital heart disease, charts of patients evaluated for cardiac disease in the neonatal period were reviewed. Comparing these to the diagnostic workup suggested by Talner and Campbell,⁴ it was possible to compile a list of pertinent findings. Historical data included prenatal conditions such as infection, hypertension or drug use, length of gestation, and other family history of congenital abnormalities. On physical exam, in addition to the vital signs, observations were made on the pulse quality (intensity in all extremities), liver size, autonomic activity and skin color. Cardiac exam included evaluation of thrills or heaves, and auscultation for the quality of the first and second heart sound, the presence of gallops or clicks and the location, timing, quality and intensity of murmurs. Laboratory evaluation included hemoglobin, hematocrit, acid-base status and blood gases in room air and with increased oxygen tension. A chest roentgenogram permitted evaluation of heart size,

pulmonary vasculature and lung fields, possible chamber enlargement, and on rare occasions, abnormalities in the location of abdominal viscera, the heart and the aorta. Electrocardiographic data included conduction patterns, chamber hypertrophy, and direction of electrical forces.

Several criteria must be met before these symptoms can be employed in a statistical model. Limited by the ability of digital computers to handle only numerical data, symptoms must be organized into discrete units. For example, a white count, elevated or normal, or plantar reflexes, up or down, are distinct possibilities. Findings which fall into subjective categories - for example, the degree of pain in arthritis - either require special handling or must be omitted. Modifications might involve definitions of the gradations. Pain, if strictly defined into categories of absent, mild, moderate or severe, represents discrete categories.

Second, if a symptom is to be included, it must have a known relationship to the diagnosis. For example, a history of maternal infection other than rubella may increase the likelihood of congenital heart disease, but may make no contribution to the specific diagnosis.

Third, definite and consistent criteria must be established for each variable. Cardiac size is difficult to assess in the neonatal period. To use such a symptom

requires establishment of specific standards.

Fourth, it is important that data be organized in the form it is obtained. For example, murmurs may be defined as "loud" or "soft" by standardized phonocardiographic deflection. If a phonocardiogram is not routinely obtained, the significance of loud and soft may be lost.

Finally, one must remember Lusted's admonition to avoid related symptoms. If arterial desaturation is used, cyanosis must be omitted. Desaturation, while not completely correlated with skin color, is both the cause of and highly related to cyanosis.

With the assistance of three experts in neonatal cardiology, Dr. Norman Talner and Dr. Marie Browne of the Department of Pediatrics, and Dr. Allan Simon of the Department of Radiology, it was possible to select the general categories of data which were utilized in the diagnosis of neonatal cardiological disease. For many of the observations, the separation between normal and abnormal was apparent. The location of the aorta, if observed, could only be left or right. A patient could, for all practical purposes, be only male or female. For other observations, particularly those employing quantification as in pulse rate, oxygen saturation, etc., it was necessary to establish cut-offs

between normal and abnormal. Upon examination of a large number of charts, it became apparent that certain groupings of data occurred within each disease and often these groupings were consistent between diseases. For example, in examining the data on the age of the patient at initial presentation, it was readily apparent that certain diseases manifested early in life - aortic atresia (hypoplastic left heart), persistent fetal pathways, and transposition of the great arteries. Others, such as aortic or pulmonary stenosis, Tetralogy of Fallot or Ebstein's abnormality, generally presented later in the newborn period. The age which separated the early from the late group was approximately three days of age: while some patients with valvular atresia presented later in life and others with stenosis presented earlier, it was possible to separate them best using the age three days.

For arterial desaturation, three groups of values were apparent: the lowest values were present in the transposition complexes, significant desaturation was present in right to left shunts, and minimal or no desaturation was present in other patients. Limited by the availability of data, it became necessary to group pO_2 values into only two groups, those with a pO_2 less than 45 and those above. In the absence of blood gas values,

the appearance of generalized cyanosis (not just perioral) was taken to correlate approximately with the lower range of arterial oxygenation, an assumption which could be inaccurate in the face of variations of hemoglobin concentrations.

Certain data, while reported consistently in the charts, was unable to satisfy the criteria established previously. For example, the "quality" of the first heart sound, while consistently noted in the charts, did not lend itself to use. The frontal axis of the p-wave, while frequently noted, was never abnormal, and consequently provided no data base.

The greatest difficulty arose in the selection of criteria for ventricular hypertrophy. Great variation exists in electrocardiographic ventricular patterns during the newborn period. Absolute criteria for hypertrophy patterns are available in the literature,⁵ but are so restrictive that patients documented to have hypertrophy may not satisfy the criteria.⁶ To surmount this difficulty, the R to S ratio in V_1 was selected as an indicator of the relative presence of ventricular forces. While patients with an increased or decreased R to S ratio may or may not be considered clinically to have hypertrophy; if the symptom-disease matrix is organized in the identical format rather than in terms of left,

or bi-ventricular hypertrophy, operation of the Bayesian system is not disturbed.

Similar choices were made for each of the other pieces of data. While controversy often arose and certain omissions were necessary - for example, the presence of Down's Syndrome may be very important in suggesting an endocardial cushion defect - the data selected included the majority considered in the clinical diagnosis. Another omission involves the presence of chamber enlargement on chest roentgenogram: a consulting radiologist with experience in neonatal angiography⁷ suggested that specific chamber enlargement cannot be ascertained accurately in the newborn period.

Echocardiographic data have also been omitted. While echos may be very valuable in determining the presence of a valve or the size of a chamber, the use of echocardiography is not sufficiently standardized, nor is there adequate data available to incorporate it into a study of this kind. Later, as sufficient data is collected, it may be possible to incorporate these findings into a Bayesian model.

To simplify the eventual computations, the symptoms were organized into the format of choices, for example, gestational age was divided into premature or term, the second heart sound as single or split and pulses as normal, increased or decreased. A patient can have only

one of these options, or if the observation was omitted, none.

Disease and Incidence Data

A number of series of congenital heart disease appear in the literature: several quantify consecutive autopsies,⁸ others involve catheterization or clinical data.⁸ Through these and review of one-hundred ninety charts at Yale New Haven Hospital, it was possible to compile the heart diseases of the newborn period. A complete listing is very extensive, involving numerous combinations of defects, and many rare abnormalities. But, it can be greatly simplified if defects are organized collectively; for example, the presence of TGA (Transposition of the Great Arteries) in association with VSD is not dissimilar from TGA with an atrial septal defect. In both cases, the patients have "Transposition Complexes" and can be grouped together under the diagnosis TGA. Patients with a VSD in association with either an ASD or PDA frequently display symptoms of their predominant shunt, usually at the ventricular level, and are symptomatically not unlike patients with a VSD alone.

The final selection of diseases (Figure 18) reflected several criteria:

1. the disease must represent a specific entity

DISEASES

1. Pulmonary atresia.
2. Pulmonary stenosis.
3. Tricuspid atresia with a VSD.
4. Tricuspid atresia without a VSD.
5. Truncus arteriosus.
6. Tetralogy of Fallot.
7. Ebstein's disease.
8. Patent ductus arteriosus.
9. Ventricular septal defect.
10. Endocardial cushion defect.
11. Transposition of the great arteries.
12. Anomalous pulmonay venous drainage.
13. Coarctation with a VSD.
14. Coarctation of the aorta.
15. Hypoplastic left heart.
16. Aortic stenosis.
17. Normal.
18. Primary pulmonary disease.
19. Persistent fetal pathways.
20. Primary myocardial disease.

FIGURE 18

and if it included two or more diseases the symptoms within each must be similar (thus, coarctation of the aorta was separated into those with and those without a VSD)

2. The incidence of the disease must be sufficient to permit the accumulation of adequate data - very rare diseases such as tri-atrial heart were omitted.
3. where combinations of diseases resulted in a new entity it had to be included separately - for example the combination of pulmonary stenosis with a ventricular defect is the Tetralogy of Fallot, an entity distinct from its components.

The list includes two non-cardiac entities. "Primary pulmonary disease" represents those patients with no intrinsic cardiac defect whose symptoms, tachypnea, tachycardia, murmurs or cyanosis, are referable to a pulmonary process. In the category "normal" are patients with neither anatomic cardiac defects nor pulmonary disease who are evaluated for arrhythmias, transient cyanosis or prematurity. Myocardial disease involves myocardopathies of metabolic origin or coronary artery abnormalities which produce myocardial ischemia.

Series from the literature were inappropriate for use as incidence data. Neither autopsy series nor catheterization data correlated exactly with the population under consideration. The most appropriate data was obtained from records maintained by the Section of Pediatric Cardiology. The "Regional Infant Care Program" log included all patients catheterized or autopsied during a several year period and for a brief period, all the patients consulted by the service, whether or not an anatomical diagnosis was ever obtained. Through analysis of these records, it was possible to determine the incidence of each of the defects over a several year period. The incidence of "normal", "primary pulmonary disease" and PDAs, three diagnoses for which patients were frequently not catheterized, was computed from the brief period when all consults were recorded. Data on the remaining disease was tabulated from the complete series of catheterization or autopsy patients.

Symptom-Disease Matrix

To complete the symptom-disease matrix, three sources were employed. First, one-hundred ninety charts of patients previously evaluated by the pediatric cardiologists were examined. Second, the literature

was reviewed, but little of data existed in the format required by a Bayesian analysis. The sex distribution - male to female ratio for each disease - varies little between sub-populations and can be abstracted from the literature.⁹ Some data on PR and QRS intervals appeared in a text on electrocardiographic data.¹⁰ Radiographic data on aortic arch location was readily available from the literature in a format applicable to this study.¹¹

The third source was the three clinicians mentioned previously. Data on history and physical findings was obtained from Dr. Talner. Electrocardiographic and auscultatory findings were contributed by Dr. Browne. Radiographic data was provided by Dr. Simon.

With these three sources of data, it was possible to perform the initial crude analyses. Reviewing the one-hundred ninety cases, and comparing each to the three sets of data, a final symptom disease matrix was compiled which appears in Appendix 3 . A casual glance at the data will reveal some apparent errors. Many of the symptom complexes total greater than 100% resulting from the absence of 0.00 as an entry anywhere in the matrix. Instead the value 0.01 or greater is employed as a "default" value for a symptom generally absent from a disease. For others, symptoms felt to be distinct clinically were combined: for example, heart size, while recorded as normal, moderately or markedly

enlarged, is dealt with statistically as normal or enlarged. For other symptoms, all the choices are given an equivalent value - .50 or fifty percent. This reflects the lack of significance of these variables. For example, the presence or absence of a third heart sound may be important to a clinical diagnosis. However, the absence of sufficient data and the difficulty of auscultation of a third heart sound eliminated it as statistically significant. In neonatal cardiology, or in any area of medical diagnosis, the differences between observers may play an important role in diagnosis. For example, the difficulty of the auscultatory exam of a tachycardic newborn may yield spurious data. If the second heart sound is reported as single the clinician is more prone to such diagnoses as valvular atresia, transposition or truncus arteriosus rather than a septal defect or patent ductus. A more careful examination may reveal the presence of two components to that sound, suggesting a different group of diagnoses.

Clinicians could minimize the extent of observer variation and its effect on diagnosis were it not for many studies¹² such as one by Yerushalmy¹³ in the area of radiology. A group of physicians were presented with roentgenograms which they were to evaluate for the presence of pulmonary tuberculosis. The interpreters were expected to read the films by their usual criteria;

none were defined. All the clinicians missed approximately twenty-five percent of the positive films. When the films were reviewed by the same observers several months later, the same reader was apt to change his diagnosis for about one of every five positive cases.

Of particular importance to neonatal cardiological diagnosis is the evaluation of pulmonary vasculature. The lung fields could demonstrate a number of patterns including decreased, normal and increased flow, and a pattern of "flow-failure" or venous obstruction. Observers generally classify the vasculature into one of these categories with such qualifications as "markedly" or "upper" or "lower limits". Different clinicians will review the films, both with and without clinical data, and make different observations. In Arnois' study¹⁴ three physicians from a Department of Cardiology and Radiology, all experienced in congenital cardiac disease, were shown 128 films on three occasions. The films were selected so that 32 demonstrated decreased and 32 increased vascularity, all documented by catheterization; the remaining 64 were normal. The clinicians were shown the masked lung fields on two occasions; on the final viewing the masking was removed. "The great majority of errors was made in differentiating between normal and decreased vasculature In contradistinction, the differentiation between normal and

increased vasculature was good. An infrequent but interesting error consisted in the scoring of decreased vascularity as increased. It is of particular interest that this error occurred only in cases of the Tetralogy of Fallot. The converse - scoring of increased vasculature as decreased - did not occur in this study. The total number of errors in the group of patients with pulmonary undervascularity was 136 out of 288 readings or 47 percent."¹⁵

These data would suggest that differentiation between normal and decreased flow, while important clinically, is a difficult radiographic observation. For the purposes of this study, no attempt has been made in the data base to employ the distinction. If a clinician observed decreased flow it is placed in the category of decreased-normal flow. Similarly, the distinction between increased flow and "flow-failure" was difficult and these categories were combined.

The evaluation of murmurs is similarly difficult. Feinstein, in the Prologue to Clinical Judgment, describes his encounter with observer variation while directing a rheumatic fever prophylaxis study:

Soon after I began this new work, while making ward rounds one day, I heard a faint, but unequivocal diastolic murmur along the left sternal border of a patient in whom no murmurs had been noted either by the resident physician

of our hospital or by the physicians who had just referred the patient to us. After I demonstrated the murmur to our resident physician, he agreed that it was there and that he had failed to recognize it.

The correction of the resident's error was a simple event - part of the ordinary daily routine of clinical activities and training. Yet, as I later thought about the event, it assumed greater significance. Since the resident physician and I might have been the only two doctors who were going to listen to this patient's heart at our institution the murmur would have been undetected had I not found it. The patient would have been discharged with the same diagnosis of "no heart disease" with which she had arrived. If the murmur persisted, it would probably be found at some later date by another auscultator. Since the murmur would never have been cited previously, however, the new auscultator might falsely conclude that the murmur had arisen from insidious scarring of an aortic valve whose damage has previously been clinically imperceptible.

The insidious development of scarring in valves that initially seemed undamaged has long been regarded as a major pathogenetic mechanism in rheumatic heart disease. I began to wonder how many patients might have developed such "scarring" as a fallacy of clinical auscultation. De novo rheumatic heart disease - absent on initial examination of a rheumatic patient and found in another examination some time later - might certainly occur by insidious scarring of a valve, but could also be "created" by iatrogenic mechanisms: the abnormal murmur might have been present on both auscultatory examinations although undetected initially; or it might have been absent on both examinations, but erroneously diagnosed as present on the second. The whole concept of insidious de novo rheumatic scarring seemed to depend on clinical auscultation of the heart, and yet auscultators could sometimes be wrong.¹⁶

In reviewing charts of patients evaluated for congenital heart disease, it was apparent that murmurs were recorded differently by different observers. Beyond

the highly subjective grading of loudness, differences involve type - holosystolic versus ejection - as well as location - upper versus lower versus entire sternal border, for example. Allowance for this type of variation required significant changes in the matrix. For example, a left upper sternal border ejection murmur is not characteristic of a PDA. In several patients in the series, this was the murmur recorded by the clinician. If the symptom-disease matrix contained the value 0.00 for the incidence of that murmur with a ductus, Bayes Law could provide a probability of zero percent for that diagnosis. To avert this situation, the "default" value, for each entry is never set at zero, but at least 0.01, unless the presence of a symptom is pathognomonic of the absence of the disease. The level of the default value will reflect the vagueness of the symptom. Murmurs are particularly difficult to evaluate in this population and ejection murmurs at the left sternal border are so frequently reported that they are of little diagnostic significance. Thus, the default value for ejection murmurs at the left sternal border must be set high, 0.05, so that their presence eliminates no diagnoses. Diastolic murmurs are less frequently recorded so that the default level has been set at two percent.

Even the evaluation of differential pulses, while more precise, must allow for some observer or disease variation. For several patients in this series, differential pulses were recorded by at least one observer but at catheterization, no cardiovascular abnormality would be demonstrated to explain the finding.

The use of default value increases the total for several symptoms to greater than 1.00. While theoretically impure, this introduces no strain to a Bayesian system. To eliminate the increased totals, one need only normalize the values for each symptom, i.e., divide each entry by the total for that symptom in that disease. In practice, the denominator of Bayes Law acts as a normalizing factor and the use of totals greater than one introduces minimal error to the final diagnosis.

Data Collection

Patient data for the testing of the Bayesian system was collected at three stages. The first group of one-hundred ninety cases were selected from consecutive listings of the Regional Infant Care Program. The second group, fifty patients, was obtained by further review of these listings. An additional thirty charts were requested for these two samples but were not included in the study. In several cases, the patient expired either prior to arrival at the hospital or before

a complete evaluation by the hospital staff. For the remainder, the charts were unobtainable. Some had been lost, while others had been signed out to hospital staff as long as six months earlier and never returned. The diagnoses of these patients were obtained from the files of the Section of Pediatric Cardiology. As the distribution of diagnoses was similar to that of the remainder of the population, it was felt that omission of these patients would not compromise the study.

The charts of the patients were abstracted into a form similar to the one shown in Appendix 1 . If available, the evaluation by the most senior cardiologist who saw the patient prior to catheterization was employed. In the few cases where the cardiologists did not enter a complete note prior to catheterization or death, the most senior housestaff's evaluation was used. If in the chart or the files maintained by the cardiologists, the electrocardiogram was observed for the pertinent findings. When absent, this information was obtained from the clinician's note. In some cases, the clinicians reported only the QRS axis and ventricular patterns so that data on PR and QRS intervals was omitted from the analysis. Approximately half of the chest roentgenograms were reviewed by the author with assistance from Dr. Simon. For those films which were unobtainable, the clinician's

observations were employed.

Greater than ninety percent of these patients had undergone catheterization so anatomical diagnoses were generally available. For those patients not catheterized, clinical diagnoses were recorded. It is significant to note that these patients represent a subgrouping of the population under consideration. Since the majority had been catheterized, the sample is skewed towards the more serious defects requiring diagnostic study, and away from the patients with PDAs or other less serious heart defects.

Computer Program

The large number of calculations required by Bayes Law would discourage the user from performing the analysis manually. Through the use of a digital computer, the several thousand logical and mathematical steps can be performed in seconds and at minimal cost. The programming for this research was performed at the Yale Computer Center. Except for some initial data manipulation employing SPSS (Statistical Package for the Social Sciences¹⁷), all programming was in the Fortran IV Language,¹⁸ an IBM package for the manipulation of numerical data. Computations were performed on the IBM 370/158 using data card input and printed output.

The cost of the program is minimal. Using the most comprehensive output format, a "run" averages \$0.80. Of this, \$0.25 is a fixed "job charge" and approximately \$0.45 is spent reading the input deck and the symptom-disease matrix. Less than \$0.10 is actually spent reading in the patient symptom information, performing the thousands of logical and mathematical computations, and printing the output. The cost of reading the deck and matrix could be decreased through the use of binary input. Even the use of an on-line, interactive program would probably amount to less than \$1.00 per day.

The program involves three steps. First, the symptom-disease matrix is read into the computer, followed by data on one or more patients. Next, the computations are performed using Bayes Law. The results - the probability of each of the twenty diseases occurring with the given set of symptoms - are then printed.

The actual program is shown in Appendix . The majority of steps are either input-output instructions or "book-keeping", i.e., instructions which are necessary only to the computer program and have no bearing on the actual Bayesian analysis. The remainder of the program consists of the instructions for the calculations of Bayes Law. As shown in Chapter III, Bayes Law, with the

assumption of conditional independence of variables,
can be written as:

$$P(D_i/S_1, \bar{S}_2, \dots, S_n) = \frac{P(D_i)P(S_1/D_i)(1-P(S_2/D_i))\dots P(S_n/D_i)}{P(D_1)P(S_1/D_1)(1-P(S_2/D_1))\dots P(S_n/D_1) + P(D_2)P(S_1/D_2)(1-P(S_2/D_2))\dots P(S_n/D_2) + \dots + P(D_m)P(S_1/D_m)(1-P(S_2/D_m))\dots P(S_n/D_m)}$$

Since the symptoms and the symptom-disease matrix were organized into choices, \bar{S}_j (symptom S_j being absent) would not occur. Instead, some other symptom, possibly S_{j+1} or S_{j-1} , would correspond to the absence of symptom S_j . Bayes Law now reverts to:

$$P(D_i/S_1, S_3, \dots, S_n) = \frac{P(D_i)P(S_1/D_i)P(S_3/D_1)\dots P(S_n/D_i)}{P(D_1)P(S_1/D_1)P(S_3/D_1)\dots P(S_n/D_1) + P(D_2)P(S_1/D_2)P(S_3/D_2)\dots P(S_n/D_2) + \dots + P(D_m)P(S_1/D_m)P(S_3/D_m)\dots P(S_n/D_m)}$$

As noted in previous sections, the matrix is in the form of probabilities, while the data on each patient or case is in the form of ones or zeroes, i.e., the patient either has or does not have the symptom. To determine the probability of each disease, one needs to calculate both the numerator and denominator. Fortunately, the denominator is identical for every disease for a given patient, and is in fact, the sum of all the numerators. A great savings

is accomplished by calculating each of the twenty numerators separately, finding their total and dividing each by the total. This yields results in the form of probabilities. To express the results as percentages, each probability is multiplied by one hundred percent.

FOOTNOTES

1. Lawrence Fitzgerald, John Overall, Clyde Williams, "A Computer Program for Diagnosis of Thyroid Disease" American Journal of Roentgenology (August, 1966, Vol. 97, No. 4) pp. 901-905.
2. C. A. Nugent, H. Warner, J. Dunn, "Probability Theory in the Diagnosis of Cushing's Syndrome" Journal of Clinical Endocrinology (Vol. 24, No. 7, 1964) pp. 621-627.
3. Norman Talner, A. G. M. Campbell, "Recognition and Management of Cardiologic Problems in the Newborn Infant" Progress in Cardiovascular Diseases (Volume 15, September-October, 1972) pp. 159-189.
4. Ibid.,
5. Warren Guntheroth, Pediatric Electrocardiography (Philadelphia, W. B. Saunders, 1966).
6. Personal communication, Dr. Marie Browne.
7. Personal communication, Dr. Allan Simon.
8. Several appear in Chapter 6 of R. Rowe, A. Mehrizi, The Neonate with Congenital Heart Disease (Philadelphia, W. B. Saunders, 1968).
9. John Keith, Richard Rowe, Peter Vlad, Heart Disease in Infancy and Childhood.
10. George Burch, Nicolas dePasquale, Electrocardiography in the Diagnosis of Congenital Heart Disease (Philadelphia, Lea and Febinger, 1967).
11. Larry Elliot, Gerold Schiebler, X-Ray Diagnosis of Congenital Cardiac Disease (Springfield, Illinois, Charles Thomas, 1968).
12. For a comprehensive review of observer error in medical diagnosis see Chapter 4 and Appendix 2 of Lee Lusted, Introduction to Medical Decision Making (Springfield, Illinois, Charles Thomas, 1968).
13. J. Yerushalmy, "Reliability of Chest Radiography in the Diagnosis of Pulmonary Lesions" American Journal of Surgery (January 1, 1955, Vol. 89) pp. 231-240.
14. Dolores-Crystal Arnois, Frederic Silverman, Malcolm Turner, "The Radiographic Evaluation of Pulmonary Vasculature in Children with Congenital Heart Disease" Radiology (May, 1959, Vol. 72) pp. 689-697.
15. Ibid., pp. 692-694.

16. Alvan Feinstein, Clinical Judgment (New York, Robert Kreiger Publishing Company, 1967) pp. 2-3.
17. Norman Nie, Dale Bent, Hadlai Hull, Statistical Package for the Social Sciences (New York, McGraw Hill, 1970).
18. International Business Machines, IBM System/360 and System/370 Fortran IV Language (New York, International Business Machines, 1974).

CHAPTER V

RESULTS AND DISCUSSION

Of the three populations studied, the first (190 patients) was employed in the construction of the symptom-disease matrix so evaluation of their results would be inappropriate. The matrix had been constructed previous to the analysis of the second group (50 patients) so a statistical evaluation, while fraught with the difficulties present in any retrospective chart review, is valuable. The third population, the twenty patients evaluated prospectively, represents the "target" population. While small in number, this is the ideal group for statistical review.

The data analysis for each patient was printed in a format similar to the one shown in Appendix 4. The clinical (pre-catheterization) diagnosis is printed along with the computer's "predicted-diagnosis" and a differential, any diagnosis which received greater than a one percent probability in the Bayesian analysis. The choice of one percent as the limit for a diagnosis to be considered significant was an arbitrary one arrived at after review of the initial one-hundred ninety patients. With knowledge of the operation of a Bayesian system, one learns that

simple shifts in the data base can produce significant changes in the predicted likelihood for any individual disease. Ignoring a diagnosis with a probability of 1.5% could limit the value of a Bayesian analysis, a program which does not "diagnose" a clinical situation, but rather predicts the likelihood of the given set of symptoms in each of the disease under consideration. Theoretically, a situation could arise in which each of the twenty diseases could enter into the differential diagnosis with a probability of greater than one percent, and the Bayesian analysis would have done little to clarify the diagnostic situation. In practice, the differential diagnosis generally approximates five diseases, on occasion it was as large as nine and in several instances only one diagnosis received greater than a one percent probability.

In reviewing the data for the original cases, an additional problem became apparent. The analysis had difficulty differentiating the categories normal and primary pulmonary disease. This reflects the similarity of data for the two categories. In addition, the majority of patients who are diagnosed as normal are evaluated for tachypnea or cyanosis, both of which, if non-cardiac, are most likely of pulmonary origin. The inability of the program to differentiate these categories has required that they be combined. The new category is entitled

"non-cardiac disease". Data provided in the analyses consists of the total of the two separate categories rather than a separate entity.

For the fifty patients studied retrospectively, the Bayesian "predicted diagnosis" agreed with the catheterization or final clinical diagnosis in fifty-two percent. For nine (18%), the clinical diagnosis received the second greatest percentage in the analysis, and for an additional three (6%), the clinical diagnosis corresponded to the third most likely computer diagnosis. In almost ninety percent of the cases, the clinical diagnosis was listed in the computer's differential diagnosis, i.e., in only twelve percent of the cases did the actual diagnosis receive less than a one percent probability. Employing a more conservative cut-off of two percent would increase the percentage of total omissions to sixteen percent and a highly conservative limit of five percent would increase it to eighteen percent. The distribution of clinical versus predicted diagnoses is shown in Figure 19. Only three of the nineteen diagnoses were predicted far more often than they occurred. Tricuspid atresia with a VSD occurred four times in this series and was predicted six times. Coarctation of the aorta occurred twice and was predicted twice as often. Persistent fetal pathways was present in only one patient but it was predicted in two others.

ACTUAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
PREDICTED																				
1. Pulmonary atresia	3		1								1									5
2. Pulmonary stenosis		1				1	1													3
3. Tricuspid atresia with VSD			1								2					1				4
4. Tricuspid atresia w/o VSD	1																			1
5. Truncus arteriosus									1											1
6. Tetralogy of Fallot											1									1
7. Ebstein's disease																				1
8. Patent ductus arteriosus	1							6		1										8
9. Ventricular septal defect						1		1	4											6
10. Endocardial cushion defect																				5
11. Transposition	1					1	1				2									4
12. Anomalous pulm. ven. drain.																				4
13. Coarctation with VSD													1	2					1	6
14. Coarctation																				1
15. Hypoplastic left heart																				1
16. Aortic stenosis						1														1
17. Non-cardiac disease						1														1
18. Persistent fetal pathways						1	1											1		3
19. Myocardial disease	6	1	2	0	0	6	3	7	5	1	6	0	1	2	6	1	1	1	1	50

FIGURE 19

Predicted versus actual diagnosis for 50 retrospective cases.

In contrast, the Tetralogy of Fallot occurred in six patients but was predicted only once. For the remainder, the frequency of clinical and computer diagnoses was similar, within plus or minus one.

For the twenty patients studied prospectively, the Bayesian predicted diagnosis agreed with the final clinical or catheterization diagnosis in sixty percent of the cases. In three cases (15%) the actual diagnosis was given the second greatest percentage by the computer program and in one (5%) the actual diagnosis entered the computer's differential as the third most likely. In only one of the twenty cases was the computer unable to give the appropriate diagnosis the one percent probability required for entry into the differential diagnosis.

This was a premature male (Patient 1, Appendix 4) with certain unusual data not considered by the program. In addition to the abnormal findings of a single second heart sound, an ejection murmur at the left lower sternal border, and left axis deviation, the chest roentgenogram revealed the presence of situs inversus with levocardia, a combination associated with specific abnormalities¹. Using the clinical data, the program yielded two likely diagnoses, a VSD (38.4%) and a patent ductus (29.8%). On reevaluation of the patient one week later (prior to catheterization) the program suggested either a VSD (35.8%)

or an abnormality of the right ventricular outflow tract (pulmonary atresia or pulmonary stenosis). In neither did the analysis yield greater than the one percent probability for the correct diagnosis - an endocardial cushion defect.

The second cushion defect in the prospective series illustrates an omission in the symptom list. The patient's clinical data, a full-term female with hepatomegaly, an ejection click, a systolic murmur at the left sternal border, a northwest electrocardiographic axis, cardiomegaly and increased pulmonary blood flow, is shown in Appendix 4, patient 2. The predicted diagnosis, a ventricular septal defect, was considered five times as likely as the actual diagnosis, an endocardial cushion defect. In addition to the two findings of increased pulmonary flow and a northwest axis which were highly suggestive to the clinicians of the actual diagnosis, the patient also displayed Down's Syndrome, making the clinicians virtually certain of their diagnosis. The program, not taking into consideration the presence of Down's Syndrome and its association with cushion defects, utilized incidence data which penalized cushion defects strongly (0.014 for ECDs and 0.113 for VSDs) and subsequently was unable to give the appropriate results.

Another error occurred in a full-term male (Patient 3) who was evaluated at birth. Clinically the patient

displayed only tachypnea, but the chest roentgenogram demonstrated borderline cardiomegaly and a venous obstructive pattern. Using the data, the program suggested a patent ductus (50%) and gave the correct diagnosis, non-cardiac disease, less than a two percent probability. Clinically this infant of a diabetic mother displayed no evidence of a ductus other than an abnormal chest roentgenogram and the x-ray findings were attributed to the maternal condition. The computer program, lacking appropriate data on infants of diabetic mothers was unable to ascertain the appropriate diagnosis.

For several other patients, the program not only provided the correct diagnosis but gave a differential diagnosis which very much resembled the clinical differential. For Patient 4, a term male with hepatomegaly, cyanosis, normal auscultatory exam and normal chest roentgenogram, the program gave a great likelihood to both non-cardiac disease (68.4%) and persistent fetal pathways (24.4%), the two major components of the clinical differential. The final diagnosis was non-cardiac disease, which agreed with the predicted diagnosis.

Another term male, Patient 5, was noted to have cyanosis, increased pulses, a single second heart sound, an ejection murmur at the left upper sternal border, a normal cardiogram and a chest roentgenogram demonstrating increased heart size and a venous obstructive pattern.

Three diagnoses - a VSD (28.6%), PDA (27.0%) and truncus arteriosus (23.0%) entered prominently into the computer's differential. The clinicians suspected a PDA because of the pulse intensity and apparently briefly considered the other two diagnoses suggested by the computer. At catheterization a truncus was demonstrated.

For another patient, the computer's diagnosis, while incorrect, was similar to the initial clinical impression. A term male, Patient 6, with decreased pulses, cyanosis, two murmurs and a normal chest roentgenogram was considered clinically to have either a myocardial disorder or persistent fetal pathways. The computer strongly suggested persistent fetal pathways (60.7%) and suggested only a slight probability of a cardiomyopathy (2.2%). At catheterization a complex defect was found which fell into the transposition complex, a diagnosis which the computer had given its second greatest probability (19.0%) but had not entered prominently into the clinical diagnosis.

Several additional cases are displayed in Appendix 4. These demonstrate both correct and incorrect predictions by the program.

Further evaluation of the results might include examination of the probabilities yielded by the computer, or to compare them to clinicians' predictions. In Warner's

study,² a method was formulated to compare clinicians' predictions to those of the computer. He created a "Diagnostic Performance Index" which is"

$$\text{DPI} = \bar{P} \times F$$

where \bar{P} is the mean probability assigned to the correct diagnosis, and F is the fraction of cases in which the correct diagnosis was given probability greater than one percent. As his clinicians organized their differential diagnosis lists in a format similar to that of the computer, including estimating probabilities, Warner was able to provide meaningful comparisons between clinical and computer diagnoses. In the present study, the clinicians did not estimate probabilities for their predicted diagnoses, and rarely discussed the relative likelihood of different elements of their differential. Thus, it is difficult to make statistical comparisons between the clinician and computer.

In addition, it is important to not place too great a value on the actual probabilities suggested by the computer program. For example, one patient in the retrospective series, a premature female (Patient 12) with normal pulses, a holosystolic murmur at the left lower sternal border, cardiomegaly and increased pulmonary

flow was felt clinically to have a ventricular septal defect. Two diagnoses, a PDA (49.7%) and a VSD (49.4%) entered the computer's differential. At catheterization, a ductus and an intact ventricular septum were demonstrated. To fault the computer for not assigning a greater probability to the appropriate diagnosis would be in error. The program responded to the data it was given by suggesting that two diagnoses - a PDA and VSD - were almost equally likely. In other cases, the low probability assigned to the appropriate diagnosis does not imply that the diagnosis is unlikely, but rather, that other diagnoses must be seriously entertained in the differential. As the number of possible diagnoses increases, the probability assigned to each must necessarily decline.

In spite of the absence of data comparing the clinicians' predictions to those of the computer, some observations can be made. For sixteen of the twenty patients (80%) the clinicians' original diagnosis agreed with their final diagnosis compared to the computer's sixty percent performance. For two patients (10%), the actual diagnosis entered into the clinician's prediction but was not considered to be the most likely diagnosis. For two patients (10%), the actual diagnosis was given a very low probability by the clinicians. In addition,

there were several instances where the clinicians, in examining the computer's complete differential, felt that it was not only more complete, but also more logical, than their own.

Through review of the twenty prospective cases, it has become apparent that there are a number of weaknesses in the data base, the symptom-disease matrix. First, there are great errors in the incidence data. The category "non-cardiac" disease was given only a seven percent incidence in the matrix but was the final diagnosis for thirty-five percent of the patients in the prospective series. This is not surprising: it was known that the retrospective series was heavily skewed towards the more seriously ill patients and included few patients with pulmonary disease or who were normal. The incidence of non-cardiac disease was estimated from data which supposedly included all patients consulted by the service, whether or not catheterization was performed. In retrospect, it is most likely that many patients with non-cardiac disease whose evaluations were minimal were not entered into the Regional Infant Care Program records. Consequently, the incidence data for non-cardiac disease is in error. Similarly, PDAs, a diagnosis which is frequently made without invasive techniques, comprised twenty-five percent of the prospective cases in contrast to 14.7% incidence in the matrix.

Two of the patients in the prospective series had endocardial cushion defects, yielding an incidence seven times that of the symptom-disease matrix. While both these patients were catheterized, it is likely that the diagnosis of a cushion defect is frequently made without catheterization. In patients with Down's Syndrome and a northwest electrocardiographic axis, the diagnosis may be readily apparent. The presence of Down's Syndrome may therefore serve as a contraindication to aggressive diagnostic study. Alternatively, in a larger series of prospective cases, no additional endocardial cushion defects might appear and the incidence data could more closely approximate that of the matrix.

The incidence of many of the remaining diagnoses did not coincide with the predictions. There was only one ventricular septal defect (5%) compared to an incidence figure of 11.3%, and two transposition complexes (10%) compared to a prediction of 18.9%. Hypoplastic left heart never entered the prospective series, while it was predicted to have a 9.9% incidence in the matrix. Several possibilities exist to explain these variations. First, the distribution of congenital heart disease may be shifting due to changing patterns of pre-natal care or abortions. Second, it is difficult to make statistically valid observations of the distribution of nineteen disease

categories in twenty cases. It is also interesting to note that historically, certain diagnoses have occurred in clusters: following a period when the diagnosis of hypoplastic left heart was never made, several patients were noted in rapid succession to have that disease. The effect of unknown variables, possibly uterine viral infections or drugs, cannot be ignored in these distribution variations.

To assess the effect of the incorrect incidence data on the program, the prospective series was reanalyzed with the incidence data omitted, i.e., giving each disease an incidence of five percent. The results were generally better than when the incidence data was employed. The percentage of correct diagnoses increased to seventy-five, one diagnosis was assigned the second greatest probability and one, the third. None of the differential diagnoses omitted the correct diagnosis. The lowest probability assigned to the correct diagnosis was 4.1%. By omitting the incidence data from the analysis, although one is violating Bayes Law, the analysis yields the comparison of the patient's symptoms to those present in each disease, ignoring the relative frequency of each disease. In two cases described previously, this variation yielded interesting results shown in Appendix 5. The predicted diagnosis for Patient 5

became truncus (77.7%) which was the catheterization diagnosis. Patient 2 was appropriately diagnosed as having an endocardial cushion defect (29.6%). For several of the patients, those with diagnoses of transposition or PDA, this variation decreased the probability assigned to the correct diagnosis, but in none of these cases did it change its location in the differential diagnosis. These observations would suggest two possibilities. First, the incidence data for the program contains serious errors which should be revised before the program can be used effectively. Second, they suggest that a variation of Bayes Law:

$$P(D/S) \propto P(S/D)$$

may have statistical significance. This formula is equivalent to the observation that a disease is more likely if the symptoms which are present resemble those likely in the disease, even if the disease is known to occur rarely. This is similar to the clinician making his diagnosis in the absence of incidence data, or employing incidence data in crude forms such as rare, frequent, etc. An error potentially introduced by this variation might be to overdiagnose rare diseases.

Other errors have also become apparent during the testing of the matrix. For example, the occurrence of decreased pulses in Ebstein's disease was set at the

default level. Subsequently, it has become apparent to the author that decreased pulses are probably common in Ebstein's disease and the value should be set higher. The data on the distribution of murmurs similarly contains errors. Ejection murmurs at the left upper sternal border were frequently reported. In many cases, the computer employed this to give a disproportionately high value to aortic stenosis. Although ejection murmurs at that location are consistent with aortic stenosis, the data must be changed to accommodate the frequency with which that murmur is reported in association with other diagnoses. Other errors in electrocardiographic data, and roentgenographic findings have become apparent and are being revised for further testing of this program. One piece of data, the occurrence of cyanosis in the "non-cardiac disease" population was surprisingly accurate. Initially, the probability was set at seventy-five and sixty percent for the "normal", "primary pulmonary disease" populations respectively. In the prospective series, of the seven patients in the "non-cardiac" category, four displayed cyanosis.

A third group of errors involves the omission of valuable entries from the symptom list. As noted, the omission of Down's Syndrome may create difficulty in the diagnosis of endocardial cushion defects. With the collection of appropriate data, it would be possible to

incorporate this finding. Other symptoms originally included in the program but ignored in the data analyses are liver size, pulse and respiratory rate, gallops and ejection sounds and the direction of T in V_1 in patients four days or older. Each was omitted because of insufficient data or the observation that data grouped too closely around the cut-off points. For example, the respiratory rate in this population is so frequently reported as seventy per minute that tachypnea appeared to be of little value as a discriminating factor. As data is collected prospectively and a larger data base is obtained, it will be possible to incorporate each of these, and possibly other symptoms and signs.

Finally, errors introduced by the assumption of independence of variables must be considered. While this assumption has not yet appeared to influence the results, errors could theoretically occur. For example, if the presence of Down's Syndrome is introduced as a variable, the incidence data is no longer independent. For patients with Down's Syndrome, there is a distinct distribution of congenital heart abnormalities,³ in contrast to the rarer occurrence of cushion defects in the remainder of the population. Findings whose simultaneous occurrence have more than a multiplicative effect on the diagnosis, would similarly defy the assump-

tion of independence. However, as a Bayesian analysis becomes unwieldy in the absence of this assumption, it is necessary despite theoretical limitations.

Despite these errors, the program has yielded results which frequently resemble the actual diagnosis. Although comparison to the clinician's predictions is difficult, the program does not appear to be significantly less accurate. With the introduction of better data it may be possible to employ the program as an effective adjunct to clinical diagnosis. The need for pediatric cardiologists would not be lessened through the use of the program. Clinicians are still needed to make the clinical observations, perform the auscultatory exam, read the roentgenograms, perform the catheterization, etc. The program is not capable of recognizing gross inconsistencies in the data. Also, the computer makes no recommendations regarding therapeutic intervention, surgery, etc. It is only capable of predicting the likelihood of diseases from a given set of symptoms.

FOOTNOTES

1. John Keith, Richard Rowe, Peter Vlad, Heart Disease in Infancy and Childhood (New York, Macmillan, 1968) pp. 557-563.
2. Homer Warner, Alan Toronto, George Veasy, "Experience with Bayes Theorem for Computer Diagnosis of Congenital Heart Disease" Annals of the New York Academy of Science (Volume 115, July 31, 1964) pp. 558-567.
3. Keith, Op. Cit., p. 808-813.

CHAPTER VI

BEYOND BAYES

In the previous two chapters, the methodology and results of a Bayesian diagnostic model have been presented. It has been demonstrated that a Bayesian model is capable, despite the recognized limitations in the data base, to assemble a differential diagnosis which is appropriate to the clinical situation, and often assigns the appropriate diagnosis the greatest probability. The utility of such a model is great and a few of the implications are to be presented in this chapter.

The value as an adjunct to clinical diagnosis is apparent. The Leeds model for the diagnosis of acute abdominal pain has demonstrated that the accuracy of a Bayesian model can exceed that of clinicians at many levels of training. Other studies by Warner and Lodwick have shown that Bayesian models can yield results similar to skilled clinicians.

The opportunity for a Bayesian model to influence a skilled clinician in his preparation of a differential diagnosis must be considered. Warner has noted that the

two clinicians involved in his studies increased their own diagnostic accuracy markedly during the study. "The extent to which the improved performance of the physicians is the result of experience in preparing data for and receiving feedback from the computer over this period of time is difficult to evaluate. It is interesting in this regard that observer (AFT), who improved the most, had the most direct contact with the computer results over the period of this study."¹ Alternatively, the improvement may have resulted from the clinician being forced to collect a complete data base.

The Leeds group has also applied their data to teaching situations.² Through the use of random number charts and a desk-top computer, artificial case histories can be created for which the student can assemble a differential diagnosis. In light of the known accuracy of this particular Bayesian model, it can be instructive to the student to compare his own diagnosis with the computers. Regretably, the accuracy of the diagnosis cannot be compared to a real patient, but "there are isolated occasions (such as when a patient scheduled for bedside teaching goes home or refuses permission) when a series of artificial substitutes might be useful. Indeed, (we) have found our series quite useful on occasion - not so much for the cases

themselves as for the subsequent discussion with the students, to whom the concept of 'certainty' in diagnosis is often new and intriguing."³

An important implication of Bayes Law is the ability to examine components of the diagnostic workup. Having compiled and tested a symptom disease matrix, it is possible to omit an individual symptom or finding from the analysis and determine the effect on diagnosis. If it is found to have no effect the utility of that symptom for the diagnosis of that disease is low. A mathematical approach to this problem proposed by Warner⁴ involves evaluation of the "information content" of a symptom for a disease. This "may be defined as the logarithm of the ratio of the probability that symptom x_1 is present or absent in any patient from this population.

$$I = \ln(P(x_1/y_1)/P(x_1))$$

Now this information can be either positive or negative since the probability of a particular symptom in a given disease may be greater or less than the incidence of that symptom in the group of diseases under study. However, if the information content is defined as the absolute value of the logarithm of this ratio, a number is obtained which is independent of the sign of the measure. The average information content (\bar{I}) of a given

symptom for a set of k diseases can be obtained from:

$$\bar{I} = \frac{1}{k} \sum_{\text{all } k} | \ln P(x_1/Y_k) - \ln P(x_1) |$$

Now since:

$$\frac{P(Y_k/x_1)}{P(Y_k)} = \frac{P(x_1/Y_k)}{P(x_1)} = e^{\bar{I}}$$

..., the term $e^{\bar{I}}$ is the average factor by which the calculated probability of a disease is changed by finding symptom x_1 to be present or absent in the patient to be diagnosed. This term then is a direct measure of the average value of that particular symptom in diagnosing this group of diseases." ⁵

Warner proposes three criteria for the inclusion of a symptom in a Bayesian model: ⁶

1. A symptom should be one whose presence or absence can be accurately recognized.
2. A symptom should be independent of other symptoms in any given disease.
3. A symptom should have an information content greater than 1.0.

Of these, the third offers the greatest opportunity in the application of Bayes Law to medical diagnosis. Applying these calculations to a clinical problem enables one to determine the information content of

each symptom: those with very low information contents could be omitted from the diagnostic workup. Warner notes that vague symptoms such as fatigue have a low information content in his congenital heart disease model, and these were omitted from subsequent analyses.⁷ In a more sophisticated model, the information content of a symptom can be compared to its cost. Those symptoms with a low information content to cost ratio could be omitted. Ultimately, Bayesian analysis offers a technique to structure a medical workup to optimize results while minimizing costs and inconvenience for the patient.

Another use of Bayes Law, in sequential processing and optimization, comes from models offered by Gorry⁸ and Ginsberg.⁹ In every clinical situation, the clinician proceeds in a sequential manner, collecting initial data, formulating a clinical impression, collecting additional data, revising the differential and so on. "The value of information obtained from a test is determined by the contribution that this information makes to improving the current view and hence reducing the risk of misdiagnosis with its associated cost. The more information the doctor obtains about the patient, on the average, the less risk of a possible misdiagnosis. Hence, the doctor is inclined to perform many tests. On the other hand, the tests available to him are not without

some cost in terms of patient discomfort, time of skilled persons, money, etc. Thus, there is a conflicting tendency to hold the number of diagnostic tests to a minimum."¹⁰

Gorry, in a reanalysis of the data collected by Warner, proposes to evaluate a patient for congenital heart disease in a sequential pattern. The clinician might observe the age of the patient and check for cyanosis and murmurs. He may next obtain an EKG or perhaps he would choose a chest roentgenogram. Gorry optimizes this decision by determining which single test is most likely to clarify the diagnostic question. An initial description of the patient is fed to the computer and a set of Bayesian probabilities determined. Then, each of the other symptoms is entered separately into the program, along with the "cost" of the test (in this case, all equivalent) and the seriousness of misdiagnosis for each disease (again, all equivalent in this case). The test which optimizes the results is requested and the calculations performed again. When the introduction of additional information no longer contributes to the certainty of the diagnosis and the "cost" of misdiagnosis is minimized, the program terminates. Using sequential processing, the program yields results statistically comparable to Warner's. Of signifi-

cance is the program's use of fewer than seven tests on the average, compared to the thirty-one required by Warner's initial program.

Ginsberg in his decision analysis model for the pleural effusion syndrome, introduces a number of variables which the clinician uses but have been omitted from earlier Bayesian models. The cost of tests, the time delays and risks involved in ordering complex tests or laboratory procedures and the value of further diagnostic intervention are all considered. At each point in the analysis, the program optimizes its decisions. If a lab test or procedure adds little to the diagnosis in light of the present clinical picture, it is not requested by the program. In contrast, any inexpensive but valuable screening procedure such as a tuberculin test, enters early into this type of model. Through the introduction of parameters of cost, risk and time delay in clinical diagnosis, despite the different values placed on these factors by different clinicians, it is possible to expand statistical decision analysis to more complex medical problems.

The role which Bayes' Law and computational devices will play in the future of medical diagnosis depends upon physician acceptance. The opportunities can be well illustrated by an analogy. "The clinician often uses a stethoscope to augment his ability to hear sounds

emanating from within a body cavity. Sometimes the clinical picture is clearcut; in this instance the clinician merely uses his stethoscope to confirm his previous assessment of the patient. Sometimes the results which the clinician obtains from the use of the stethoscope are difficult to interpret or are at odds with what the clinician 'feels' about the case - in such circumstances the clinician is at liberty to disregard the evidence from his 'machine'. But in a proportion of cases the evidence the clinician obtains will alter his impression of the case sufficiently to make him seek additional evidence and this in turn will lead him to the correct diagnosis."¹¹

So it is with a Bayesian program. By helping to sort through large quantities of data, an area in which physicians are known to be relatively weak¹², the computer can provide suggestions which may aid the clinician in arriving at the appropriate diagnosis.

FOOTNOTES

1. Homer Warner, Alan Toronto, George Veasy, "Experience with Bayes' Theorem for Computer Diagnosis of Congenital Heart Disease" Annals of the New York Academy of Science (July 31, 1964, Vol. 115, pp. 566-567.
2. F. T. deDombal, Jane Horrocks, et al., "Production of Artificial 'Case Histories' by Using a Small Computer" British Medical Journal (June 5, 1971, Vol. 2) pp. 576-581.
3. Ibid., p. 580.
4. Homer Warner, "Medical Diagnosis Using a Digital Computer" Proceedings of the Third IBM Medical Symposium (1963) pp. 305-317.
5. Ibid., p. 307.
6. Loc. Cit.
7. Alan Toronto, George Veasy, Homer Warner, "Evaluation of a Computer Program for Diagnosis of Congenital Heart Disease" Progress in Cardiovascular Diseases (January, 1963, Vol. 5, No. 4) p. 374.
8. Anthony Gorry, Octo Barnett, "Sequential Diagnosis by Computer" Journal of the American Medical Association (September 16, 1968, Vol. 205, No. 12) pp. 849-854.
9. A. S. Ginsberg, Decision Analysis in Clinical Patient Management with an Application to the Pleural Effusion Syndrome (Santa Monica, Rand Corporation, 1971).
10. Gorry, Op. Cit., p. 850.
11. Jane Horrocks, A. P. McCann. "Computer-Aided Diagnosis: Description of an Adaptable System and Operational Experience with 2034 Cases" British Medical Journal (April 1, 1972, Vol. 2) p. 9.
12. F. T. deDombal, J. C. Horrocks, et al., "Pattern Recognition: A Comparison of the Performance of Clinicians and Non-Clinicians with a Note on the Performance of a Computer Based System" Methods of Information in Medicine (Vol. 11, No. 1, 1972) pp. 32-37.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Arnois, Dolores-Crystal; Siverman, Frederic; Turner, Malcolm, "The Radiographic Evaluation of Pulmonary Vasculature in Children with Congenital Heart Disease" Radiology (Volume 72, May, 1959) pp. 689-697.
- Barnes, Lewis; Tunnessen, Walter, et al., "Computer Assisted Diagnosis in Pediatrics" American Journal of Diseases in Children (Volume 127, June, 1974) pp. 852-858.
- Barnoon, Schlomo; Wolfe, Harvey, Measuring the Effectiveness of Medical Decisions (Springfield, Illinois, Charles C. Thomas, 1972).
- Bayes, T., "Essary towards solving a problem in the doctrin of chances", Philosophical Transactions of the Royal Society (Volume 53, 1763) pp. 370-418 (Reprinted in Biometrika (Volume 45, 1958) pp. 293-315).
- Bean, William, "The Natural History of Error" Archives of Internal Medicine (Volume 105, February, 1960) pp. 184-193.
- Best, William, "The Potential Role of Computers in Medical Practice" Journal of the American Medical Association (Volume 162, No. 10, December 8, 1962) pp. 124-130.
- Brand, Don; Dove, Henry; Meyers, Alan; "A Technique for Analyzing Clinical Data to Provide Patient Management Guidelines: A Study of Meningitis in Children" Submitted for publication.
- Brunk, H.D.; Lehr, J.L., "An Improved Bayes' Method for Computer Diagnosis" Proceedings of Conference on the Use of Computers in Radiology (University of Chicago, Center for Continuing Education, October 22, 1966).
- Burch, George; dePasquale, Nicholas, Electrocardiography in the Diagnosis of Congenital Heart Disease (Philadelphia, Lea and Febinger, 1967).
- Caceres, Cesar; Rikli, Arthur, Diagnostic Computers (Springfield, Illinois, Charles Thomas, 1969).
- Card, W. I., "The Diagnostic Process" Journal of the Royal College of Physicians of London (Volume 4, No. 3, April, 1970) pp. 183-187.
- Casey, R.S.; Perry, J.W. Punched Cards, Their Application to Sciences and Industry (New York, Reinhold, 1951).
- Cobb, Sidney, "On the Development of Diagnostic Criteria" Arthritis and Rheumatism (Volume 13, February, 1960) pp. 91-95.
- Collen, Morris; Rubin, Leonard; Neyman, Jerzy, et al., "Automated Multiphasic Screening and Diagnosis" American Journal of Public Health (Volume 54, No. 5, May, 1964) pp. 741-750.

- Committee of the American Rheumatism Association, "1958 Revision of Diagnostic Criteria for Rheumatoid Arthritis" Arthritis and Rheumatism (Volume 2, 1959) pp. 16-20.
- Cornfeld, Jerome, "Bayes Theorem" Proceedings of the Sixth IBM Medical Symposium (1964) pp. 163-196.
- Crooks, J., Murray, I., Wayne, E., "Statistical Methods Applied to the Clinical Diagnosis of Thyrotoxicosis" Quarterly Journal of Medicine (Volume 28, No. 110, April, 1959) pp. 211-234.
- deDombal, F. T., "Surgical Diagnosis Assisted by a Computer" Proceedings of the Royal Society of London (Volume 184, 1973) pp. 433-440.
- deDombal, F. T., Horrocks, J. C., et al., "Pattern Recognition: A Comparison of the Performance of Clinicians and Non-Clinicians with a Note on the Performance of a Computer Based System" Methods of Information in Medicine (Volume 11, No. 1, 1972) pp. 32-37.
- deDombal, F. T., Horrocks, J. C., et al., "Production of Artificial 'Case Histories' by Using a Small Computer" British Medical Journal (Volume 2, June 5, 1971) pp. 578-581.
- deDombal, F. T., Horrocks, J. C., et al., "Simulation of Clinical Diagnosis: A Comparative Study" British Medical Journal (Volume 2, June 5, 1971) pp. 575-577.
- deDombal, F.T., Leaper, D.J., et al., "Computer-aided Diagnosis of Acute Abdominal Pain" British Medical Journal (Volume 2, April 1, 1972) pp. 9-13.
- Elliot, Larry; Schiebler, Gerald, X-Ray Diagnosis of Congenital Cardiac Disease (Springfield, Illinois, Charles Thomas, 1968).
- Engle, Ralph, "Medical Diagnosis: Past, Present, and Future" Archives of Internal Medicine (Volume 112, October, 1963) pp. 512-542.
- Feinstein, Alvan, Clinical Judgment (Huntington, New York, Robert Kreiger Publishing Company, 1967).
- Feinstein, Alvan; Rubinstein, Joel; Ramshaw, Walter, "Estimating Prognosis with the Aid of a Conventional Mode Computer Program" Annals of Internal Medicine (Volume 76, No. 6, June, 1972) pp. 911-921.
- Fitzgerald, Lawrence; Overall, John; Williams, Clyde, "A Computer Program for Diagnosis of Thyroid Disease" American Journal of Roentgenology (Volume 97, No. 4, August, 1966) pp. 901-905.
- Gill, P.W., Leaper, D.J., et al., "Observer Variation in Clinical Diagnosis - A Computer-Aided Assessment of its Magnitude and Importance in 552 Patients with Abdominal Pain" Methods of Information in Medicine (Volume 12, No. 2, 1973) pp. 108-113.
- Ginsberg, A.S., Decision Analysis in Clinical Patient Management with an Application to the Pleural Effusion Syndrome

(Santa Monica, Rand Corporation, 1971).

- Gleser, Malcolm; Collen, Morris, "Towards Automated Medical Decisions" Computers and Biomedical Research (Volume 5, 1972) pp. 180-189.
- Gorry, Anthony; Barnett, Octo, "Sequential Diagnosis by Computers" Journal of the American Medical Association (Volume 205, No. 12, September 16, 1968) pp. 849-854.
- Guntheroth, Warren, Pediatric Electrocardiography (Philadelphia, W. B. Saunders, 1966).
- Gustafson, John, "The Computer - For Use in Private Practice" Proceedings of the Fifth IBM Medical Symposium (1963) pp. 99-111.
- Gustafson, David; Kestly, John; Ludke, Robert; Larson, Frank, "Probabilistic Information Processing: Implementation of a Semi-PIP Diagnostic System" Computers and Biomedical Research (Volume 6, 1973) pp. 355-370.
- Horrocks, J. C., McCann, A. P., "Computer-Aided Diagnosis: Description of an Adaptable System and Operational Experience with 2034 Cases" British Medical Journal (Volume 2, April 1, 1972) pp. 5-9.
- Horvath, William, "The Effect of Physician Bias in Medical Diagnosis" Behavioral Science (Volume 9, 1964) pp. 334-340.
- International Business Machines, IBM System/360 and System/370 Fortran IV Language (New York, International Business Machines, 1974).
- Jacquez, J.A., editor, The Diagnostic Process (Ann Arbor, Malloy Lithographing, Inc., 1964).
- Johnson, Paul; Moller, James; Bass, George, "Analysis of Expert Diagnosis of a Computer Simulation of Congenital Heart Disease" Journal of Medical Education (Volume 50, May, 1975) pp. 466-470.
- Keith, John; Rowe, Richard; Vlad, Peter, Heart Disease in Infancy and Childhood (New York, Macmillan, 1968).
- Kleinmuntz, Benjamin, "Diagnostic Problem Solving by Computer" Japanese Psychological Research (Volume 7, No. 4, 1965) pp. 189-194.
- Krovetz, Jerome; Gessner, Ira; Schiebler, Gerold, Handbook of Pediatric Cardiology (New York, Hoeber Medical Division, Harper and Row, 1969).
- Leaper, D.J., Horrocks, J.C., et al., "Computer-assisted Diagnosis of Abdominal Pain Using Estimates Provided by Clinicians" British Medical Journal (Volume 4, November 11, 1972) pp. 350-354.
- Ledley, Robert; Lusted, Lee, "Reasoning Foundations of Medical Diagnosis" Science (Volume 130, No. 3366, July 3, 1959) pp. 9-21.

- Lipkin, Mack, The Case of Patients (New York, Oxford University Press, 1974).
- Lipkin, Martin, "The Likelihood Concept of Differential Diagnosis" Perspectives in Medicine and Biology (Summer, 1964) pp. 485-497.
- Lipkin, Martin; Hardy, James, "Differential Diagnosis of Hematological Diseases Aided by Mechanical Correlation of Data" Science (Volume 125, March 22, 1957) pp. 551-552.
- Lipkin, Martin; Hardy, James, "Mechanical Correlation of Data in Differential Diagnosis of Hematological Diseases" Journal of the American Medical Association (Volume 166, No. 2, January 11, 1958) pp. 113-125.
- Lodwick, G.S., "Computer-Aided Diagnosis in Radiology. A Research Plan" Investigative Radiology (Volume 1, Jan., Feb., 1966) pp. 72-80.
- Lodwick, Gwilym; Haun, Cosmo; Smith, Walter; Keller, Roy; Robertson, Eddie, "Computer Diagnosis of Primary Bone Tumors" Radiology (Volume 80, February, 1963) pp. 273-275.
- Lodwick, Gwilym; Turner, Arch; Lusted, Lee; Templeton, Arch, "Computer Aided Diagnosis of Radiographic Images" Journal of Chronic Diseases (Volume 19, No. 4, April, 1966) pp. 485-496.
- Lodwick, S.G., Turner, A.H. Computer Aided Diagnosis of Primary Bone Tumors (A brochure, December, 1965).
- Lusted, Lee, Introduction to Medical Decision Making, (Springfield, Illinois, Charles Thomas, 1968).
- Lusted, Lee; Ledley, Robert, "Mathematical Models in Medical Diagnosis" Journal of Medical Education (Volume 35, No. 3, March, 1960) pp. 214-221.
- Mount, James; Evans, John, "Computer-Aided Diagnosis - A Simulation Study" Proceedings of the Fifth IBM Medical Symposium (1963) pp. 113-127.
- Nadas, Alexander, Pediatric Cardiology, (Philadelphia, W. B. Saunders, 1964).
- Nash, F. A., "Differential Diagnosis" The Lancet (Volume 22, No. 6816, April 17, 1954) pp. 874-875.
- Neal, William, et al., "Operative Repair of Atrial Septal Defect without Cardiac Catheterization" Journal of Pediatrics (Volume 86, No. 2, February, 1975) pp. 189-193.
- Neyman, Jerzy, "Outline of Statistical Treatments of the Problem of Diagnosis" Public Health Reports (Volume 62, October 3, 1947) pp. 1449-1456.
- Nie, Norman; Bent, Dale; Hull, Hadlai, Statistical Package for the Social Sciences (New York, McGraw Hill, 1970).
- Nugent, C.A., Warner, H., Dunn, J., "Probability Theory in the Diagnosis of Cushing's Syndrome" Journal of Clinical Endocrinology (Volume 24, No. 7, 1964) pp. 621-627.

- Overall, James; Williams, Clyde, "Models for Medical Diagnosis" Behavioral Science (Volume 6, April, 1961) pp. 134-141.
- Peckham, Robert, "Betting Odds in Medical Diagnosis" American Journal of Medical Science (Volume 253, January, 1967) pp. 65-67.
- Ropes, Marian; Granville, Bennett; Cobb, Sidney, et al., "Proposed Diagnostic Criteria for Rheumatoid Arthritis" Annals of Rheumatic Disease (Volume 16, March 25, 1957).
- Rowe, Richard; Mehrizi, A., The Neonate with Congenital Heart Disease (Philadelphia, W. B. Saunders, 1968).
- Shane, Louis; Marler, Eric; Bauman, William, "Computer Simulation of Medical Logic in Clinical Decision Support System" New York State Journal of Medicine (June 15, 1972) pp. 1636-1644.
- Shepard, J., "Correspondence" British Medical Journal (Volume 2, 1972) p. 347.
- Staniland, R; Ditchburn, J., deDombal, F.T., "Clinical Presentation of Acute Abdomen: Study of 600 Patients" British Medical Journal (Volume 3, August 12, 1972) pp. 393-398.
- Swender, Phillip; Tunnessen, Walter; Oski, Frank, "Computer Assisted Diagnosis" American Journal of Diseases of Children (Volume 127, June, 1974) pp. 859-861.
- Talner, Norman; Campbell, A.G.N., "Recognition and Management of Cardiologic Problems in the Newborn Infant" Progress in Cardiovascular Disease (Volume 15, September, October, 1972) pp. 159-189.
- Taylor, T. R., "Computer guided diagnosis" Journal of the Royal College of Physicians, London (Volume 4, No. 3, April, 1970) pp. 180-202.
- Templeton, Arch; Lehr, James; Simmons, Charles, "The Computer Evaluation and Diagnosis of Congenital Heart Disease, Using Roentgenographic Findings" Radiology (Volume 87, October, 1966) pp. 658-670.
- Toronto, Alan; Veasy, George; Warner, Homer, "Evaluation of a Computer Program for Diagnosis of Congenital Heart Disease" Progress in Cardiovascular Disease (Volume 5, No. 4, January, 1963) pp. 362-376.
- Warner, Homer, "Medical Diagnosis Using a Digital Computer" Proceedings of the Third IBM Medical Symposium (1963) pp. 305-317.
- Warner, Homer; Toronto, Alan; Veasy, George; Stephenson, Robert, "A Mathematical Approach to Medical Diagnosis" Journal of the American Medical Association (Volume 177, No. 3, July 22, 1961) pp. 171-183.

- Warner, Homer; Toronto, Alan; Veasy, George, "Experience with Bayes Theorem for Computer Diagnosis of Congenital Heart Disease" Annals of the New York Academy of Sciences (Volume 115, July 31, 1964) pp. 558-567.
- Woodbury, Max, "Inapplicability of Bayes' Theorem to Medical Diagnosis" Proceedings of the Fifth International Conference on Medical Electronics (July 22, 1963) pp. 860-868.
- Yarnall, Stephen "Mathematical Models for Prognosis in Rheumatic Heart Disease" Proceedings of the Seventh IBM Medical Symposium (1965) pp. 153-186.
- Yerushalmy, J., "Reliability of Chest Radiography in the Diagnosis of Pulmonary Lesions" American Journal of Surgery (Volume 89, January, 1955) pp. 231-240.
- Ziegler, Robert, Cardiac Evaluation in Normal Infants (Saint Louis, C. V. Mosby, 1965).

APPENDIX 1 : Data Collection Form

CARD #1

UNIT #	-----	(1-6)
GESTATIONAL AGE	37-44 WEEKS	(9)
	20-36 WEEKS	(10)
AGE AT TIME OF EVALUATION	0-3 DAYS	(11)
	4-42 DAYS	(12)
SEX	MALE	(13)
	FEMALE	(14)
PULSE RATE	170-250	(15)
	50 -169	(16)
PULSE INTENSITY	NORMAL	(17)
	INCREASED	(18)
	DECREASED	(19)
DIFFERENTIAL PULSES	ABSENT	(20)
	PRESENT	(21)
LIVER-CM BELOW COSTAL MARGIN	0-2	(22)
	3-10	(23)
DIFFERENTIAL CYANOSIS	ABSENT	(24)
	PRESENT	(25)
RESPIRATORY RATE	0 -70	(26)
	71-100	(27)
PO2	0-44	(28)
	45-100	(29)
(IF BLOOD GASES ARE NOT AVAILABLE, USE GENERALIZED CYANOSIS AS EQUIVALENT TO PO2 LESS THAN 45)		
PO2 INCREASE ON 100% O2	0 -29	(30)
	30-200	(31)
PH	7.00-7.29	(32)
	7.30-7.60	(33)
S2	SPLIT	(34)
	SINGLE	(35)
S3	ABSENT	(36)
	PRESENT	(37)
S4	ABSENT	(38)
	PRESENT	(39)
EJECTION SOUNDS	ABSENT	(40)
	PRESENT	(41)
NO SYSTOLIC MURMER		(42)
NO DIASTOLIC MURMER		(43)
APEX	SYSTOLIC	(44)
	HOLOSYSTOLIC	(45)
	EJECTION	(46)
	DIASTOLIC	(47)
	CONTINUOUS	(48)
LEFT LOWER STERNAL BORDER	SYSTOLIC	(49)
	HOLOSYSTOLIC	(50)
	EJECTION	(51)
	DIASTOLIC	(52)
	CONTINUOUS	(53)
LEFT UPPER STERNAL BORDER	SYSTOLIC	(54)
	HOLOSYSTOLIC	(55)
	EJECTION	(56)
	DIASTOLIC	(57)
	CONTINUOUS	(58)
RIGHT UPPER STERNAL BORDER	SYSTOLIC	(59)
	DIASTOLIC	(60)

CARD #2

PR INTERVAL	0.04-0.11	(8)
	0.12-0.20	(9)
QRS INTERVAL	0.03-0.06	(10)
	0.07-0.15	(11)
QRS AXIS	270-300 OR 0-59	(12)
	60-180	(13)
	181-269	(14)
R/S IN V1	GREATER THAN 10	(15)
	BETWEEN 0.8 AND 10	(16)
	LESS THAN 0.8	(17)
T IN V1	POSITIVE	(18)
	ZERO OR NEGATIVE	(19)
(OMIT IF PATIENT LESS THAN FOUR DAYS OF AGE)		
PULMONARY BLOOD FLOW	NORMAL	(20)
	DECREASED	(21)
	INCREASED W/O VEN. OBST.	(22)
	INCREASED WITH VEN. OBST.	(23)
HEART SIZE	NORMAL	(24)
	MODERATELY INCREASED	(25)
	MARKEDLY INCREASED	(26)
AORTIC ARCH	LEFT	(27)
	RIGHT	(28)

CARD #3

CLINICIAN	--	(8-9)
	01 - TALNER	
	02 - BROWNE	
	03 - NUDEL	
	04 - BERMAN	
	05 - HELLENBRAND	
	06 - GLANZ	
	07 - RESIDENT	
	08 -	
	09 -	
	10 -	
DATE	-- / -- / --	(11-18)
	MONTH/DATE/YEAR	
CLINICAL DIAGNOSIS	--	(20-21)
CATH DIAGNOSIS	--	(23-24)

- 1- PULMONARY ATRESIA
- 2- PULMONARY STENOSIS
- 3- TRICUSPID ATRESIA WITH VSD
- 4- TRICUSPID ATRESIA W/O VSD
- 5- TRUNCUS ARTERIOSUS
- 6- TETRALOGY OF FALLOT
- 7- EBSTEINS DISEASE
- 8- PATENT DUCTUS ARTERIOSUS
- 9- VENTRICULAR SEPTAL DEFECT
- 10- ENDOCARDIAL CUSHION DEFECT
- 11- TRANSPOSITION
- 12- ANOM. PULM. VEN. DRAINAGE
- 13- COARCT WITH VSD
- 14- COARCTATION
- 15- HYPOPLASTIC LEFT HEART
- 16- AORTIC STENOSIS
- 17- NORMAL
- 18- PRIMARY PULMONARY DISEASE
- 19- PERSISTENT FETAL PATHWAYS
- 20- PRIMARY MYOCARDIAL DISEASE

APPENDIX 2 : Computer Program


```

0001      DIMENSION ARFAY(74,20),CASE(74),BAYES(20),PROB(20),DOC(3,10),
          *DESC(15,74),DIS(11,21),NAME(5)
0002      READ (5,100) (ARFAY(I,J),J=1,20)
0003      100 FORMAT(7X,20F3.3)
0004      DO 99 I=2,74
0005      99 READ(5,98)(ARFAY(I,J),J=1,20)
0006      98 FORMAT(7X,20F3.2)
0007      READ(5,150) DOC
0008      150 FORMAT(5X,3A4)
0009      READ(5,151) *DESC
0010      151 FORMAT(5X,15A4)
0011      READ(5,152) DIS
0012      152 FORMAT(5X,11A4)
0013      114 READ(5,101,END=509) LUNIT,CASE,NA,MON,DATE,YEAR,NLDX,NTDX,NAME
0014      101 FORMAT(16,1X,53F1.0/7X,21F1.0/7X,12,3(1X,A2),2(1X,I2),1X,5A4)
0015      IF (NLDX.EQ.0) NLDX=21
0016      IF (NTDX.EQ.0) NTDX=21
0017      WRITE(6,500) NAME,LUNIT,(DOC(I,NA),I=1,3),MON,DATE,YEAR
0018      500 FORMAT('1'/'/' PATIENT: '1,5A4'/' UNIT #: '1,16'/' CLINICIAN: DR.
          * '1,3A4'/' DATE: '1,A2,'/'1,A2,'/'1,A2'/' SYMPTOMS:'/' )
0019      DO 501 I=2,74
0020      IF (CASE(I).EQ.0) GO TO 501
0021      WRITE(6,502) (DESC(K,I),K=1,15)
0022      502 FORMAT(5X,15A4)
0023      501 CONTINUE
0024      WRITE(6,701) (DIS(L,NLDX),L=1,11)
0025      701 FORMAT('/'/' CLINICAL DIAGNOSIS:'/'/5X,11A4)
0026      WRITE(6,503)
0027      503 FORMAT('/'/' PREDICTED DIAGNOSIS:'/' )
C IF VARIABLES ARE TO BE OMITTED, CASE( )=0 MUST BE DECLARED NON.
0028      CASE(1)=1.0
0029      CASE(19)=0.0
0030      CASE(20)=0.0
0031      CASE(29)=0.0
0032      CASE(30)=0.0
0033      CASE(33)=0.0
0034      CASE(34)=0.0
0035      CASE(64)=0.0
0036      CASE(65)=0.0
0037      TOTAL=0.0
0038      DO 103 J=1,20
0039      BAYES(J)=1.0
0040      DO 102 I=1,74
0041      IF (CASE(I).EQ.0.0) GO TO 102
0042      BAYES(J)=BAYES(J)+ARFAY(I,J)
0043      102 CONTINUE
0044      103 TOTAL=TOTAL+BAYES(J)
0045      DO 104 J=1,20
0046      104 PROB(J)=100*BAYES(J)/TOTAL
0047      PROB(17)=PROB(17)+PROB(18)
0048      PROB(18)=0.0
0049      K=1
0050      DUMMY=PROB(1)
0051      DO 400 I=2,20
0052      IF (PROB(I).LE.DUMMY) GO TO 400
0053      DUMMY=PROB(I)
0054      K=I
0055      400 CONTINUE
0056      WRITE(6,401) (DIS(L,K),L=1,11),PROB(K)
0057      401 FORMAT(5X,11A4,10X,'(,F4.1,%)')
0058      PROB(K)=0.0
0059      WRITE(6,504)
0060      504 FORMAT('/'/' OTHER POSSIBLE DIAGNOSES:'/' )
0061      DO 508 M=1,19
0062      K=1
0063      DUMMY=PROB(1)
0064      DO 450 I=2,20
0065      IF (PROB(I).LE.DUMMY) GO TO 450
0066      DUMMY=PROB(I)
0067      K=I
0068      450 CONTINUE
0069      IF (PROB(K).LT.1.0) GO TO 505
0070      WRITE(6,401) (DIS(L,K),L=1,11),PROB(K)
0071      508 PROB(K)=0.0
0072      M=M+1
0073      505 IF (M.NE.1) GO TO 506
0074      WRITE(6,507)
0075      507 FORMAT(' NONE')
0076      506 GO TO 114
0077      509 CONTINUE
0078      END

```


APPENDIX 3 : Symptom-Disease Matrix

APPENDIX 4 : Sample Patients

PATIENT: 1

CLINICIAN: DR. RESIDENT

DATE: 11/ 6/75

SYMPTOMS:

LE 36 WEEKS GESTATION
LE 3 DAYS OLD
MALE
PULSE RATE GT 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE LE 70
CYANOSIS
PH LE 7.29
S2 SINGLE
S3 ABSENT
S4 ABSENT
EJECTION CLICK PRESENT
NO DIASTOLIC MURMUR
SYSTOLIC MURMUR AT THE LEFT LOWER STERNAL BORDER
PP INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 270-360 OR 0-59
R/S IN V1 LE 10 AND GE 0.8
INCREASED PULMONARY BLOOD FLOW W/O VENOUS OBSTRUCTION
MODERATELY INCREASED HEART SIZE

CLINICAL DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT

FINAL DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT

PREDICTED DIAGNOSIS:

VENTRICULAR SEPTAL DEFECT (38.4%)

OTHER POSSIBLE DIAGNOSES:

PATENT DUCTUS ARTERIOSUS (29.8%)
TRANSPOSITION OF THE GREAT ARTERIES (14.0%)
TRICUSPID ATRESIA WITH VSD (7.9%)
COARCTATION WITH A VENT SEPTAL DEFECT (2.8%)
HYPOPLASTIC LEFT HEART (2.4%)
PULMONARY ATRESIA (1.6%)
TRICUSPID ATRESIA WITH INTACT VENT. SEPTUM (1.5%)

PATIENT: 2

CLINICIAN: DR. GLANZ

DATE: 10/26/75

SYMPTOMS:

GE 37 WEEKS GESTATION
GE 4 DAYS OLD
FEMALE
PULSE RATE GT 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER GT 2 CM BELCW CM
RESPIRATORY RATE GT 70
NO CYANOSIS
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK PRESENT
NO DIASTOLIC MURMER
SYSTOLIC MURMER AT THE LEFT LOWER STERNAL BORDER
SYSTOLIC MURMER AT THE LEFT UPPER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 181-269
R/S IN V1 LE 10 AND GE 0.8
T IN V1 GT 0 IN PATIENT AGE GT 3 DAYS
INCREASED PULMONARY BLOOD FLOW W/O VENOUS OBSTRUCTION
MODERATELY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT

FINAL DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT

PREDICTED DIAGNOSIS:

VENTRICULAR SEPTAL DEFECT (51.6%)

OTHER POSSIBLE DIAGNOSES:

TOTAL ANOMOLOUS PULM VENOUS DRAINAGE	(12.7%)
PATENT DUCTUS ARTERIOSUS	(10.0%)
ENDOCARDIAL CUSHION DEFECT	(9.3%)
PULMONARY STENOSIS	(6.7%)
COARCTATION WITH A VENT SEPTAL DEFECT	(3.5%)
AORTIC STENOSIS	(2.4%)
TRANSPOSITION OF THE GREAT ARTERIES	(1.4%)

PATIENT: 3

CLINICIAN: DR. MATISOFF

DATE: 15/16/75

SYMPTOMS:

GE 37 WEEKS GESTATION
LE 3 DAYS OLD
FEMALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE GT 70
NO CYANOSIS
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO SYSTOLIC MURMER
NO DIASTOLIC MURMER
PP INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
INCREASED PULMONARY BLOOD FLOW WITH VENOUS OBSTRUCTION
MODERATELY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

NON-CARDIAC DISEASE

FINAL DIAGNOSIS:

NON-CARDIAC DISEASE

PREDICTED DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS (50.0%)

OTHER POSSIBLE DIAGNOSES:

VENTRICULAR SEPTAL DEFECT (21.9%)
TOTAL ANOMALOUS PULM VENOUS DRAINAGE (12.0%)
PRIMARY MYOCARDIAL DISEASE (8.7%)
TRANSPOSITION OF THE GREAT ARTERIES (3.6%)
NON-CARDIAC DISEASE (1.5%)
COARCTATION OF THE AORTA (1.3%)

PATIENT: 4
CLINICIAN: DR. RESIDENT
DATE: 10/23/75
SYMPTOMS:

GE 37 WEEKS GESTATION
LE 3 DAYS OLD
MALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER GT 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE GT 70
CYANOSIS
PO2 INCREASE WITH 100% O2 GE 30
PH LE 7.29
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO SYSTOLIC MURMUR
NO DIASTOLIC MURMUR
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
NORMAL PULMONARY BLOOD FLOW
NORMAL HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

NON-CARDIAC DISEASE

FINAL DIAGNOSIS:

NON-CARDIAC DISEASE

PREDICTED DIAGNOSIS:

NON-CARDIAC DISEASE

(68.4%)

OTHER POSSIBLE DIAGNOSES:

PERSISTENT FETAL PATHWAYS
COARCTATION OF THE AORTA
PRIMARY MYOCARDIAL DISEASE

(24.4%)

(4.1%)

(1.1%)

PATIENT: 5

CLINICIAN: DR. TALNER

DATE: 11/25/75

SYMPTOMS:

GE 37 WEEKS GESTATION
GE 4 DAYS OLD
MALE
PULSE RATE LE 170
INCREASED PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE GT 70
CYANOSIS
PO2 INCREASE WITH 100% O2 GE 30
PH GE 7.30
S2 SINGLE
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO DIASTOLIC MURMER
EJECTION MURMER AT THE LEFT UPPER STERNAL BORDER
PP INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
T IN V1 LE 0 IN PATIENT AGE GT 3 DAYS
INCREASED PULMONARY BLOOD FLOW WITH VENOUS OBSTRUCTION
MARKEDLY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

FINAL DIAGNOSIS:

TRUNCUS ARTERIOSUS

PREDICTED DIAGNOSIS:

VENTRICULAR SEPTAL DEFECT

(28.6%)

OTHER POSSIBLE DIAGNOSES:

PATENT DUCTUS ARTERIOSUS

(27.0%)

TRUNCUS ARTERIOSUS

(23.6%)

TRANSPOSITION OF THE GREAT ARTERIES

(10.7%)

COARCTATION OF THE AORTA

(6.6%)

HYPOPLASTIC LEFT HEART

(1.5%)

PATIENT: 6
CLINICIAN: DR. NUDEL
DATE: 12/24/75
SYMPTOMS:

GE 37 WEEKS GESTATION
LE 3 DAYS OLD
MALE
PULSE RATE LE 170
DECREASED PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE LE 70
CYANOSIS
PO2 INCREASE WITH 100% O2 LT 30
S2 SPLIT
S3 PRESENT
S4 ABSENT
EJECTION CLICK ABSENT
NO DIASTOLIC MURMER
HOLOSYSTOLIC MURMER AT THE LEFT LOWER STERNAL BORDER
EJECTION MURMER AT THE LEFT UPPER STERNAL BORDER
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
NORMAL PULMONARY BLOOD FLOW
NORMAL HEART SIZE
LEFT ACRTIC ARCH

CLINICAL DIAGNOSIS:

PRIMARY MYOCARDIAL DISEASE

FINAL DIAGNOSIS:

TRANSPOSITION OF THE GREAT ARTERIES

PREDICTED DIAGNOSIS:

PERSISTENT FETAL PATHWAYS (60.7%)

OTHER POSSIBLE DIAGNOSES:

TRANSPOSITION OF THE GREAT ARTERIES, (19.0%)
NON-CARDIAC DISEASE (10.4%)
HYPOPLASTIC LEFT HEART (2.5%)
PRIMARY MYOCARDIAL DISEASE (2.2%)
PULMCPARY STENOSIS (1.8%)
VENTRICULAR SEPTAL DEFECT (1.4%)

PATIENT: 7

CLINICIAN: DR. MATISCOFF

DATE: 12/21/75

SYMPTOMS:

LE 36 WEEKS GESTATION
GE 4 DAYS OLD
FEMALE
PULSE RATE LE 170
INCREASED PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
RESPIRATORY RATE LE 70
NO CYANOSIS
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
CONTINUOUS MURMUR AT THE LEFT LOWER STERNAL BORDER
CONTINUOUS MURMUR AT THE LEFT UPPER STERNAL BORDER
PR INTERVAL GT 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
NORMAL PULMONARY BLOOD FLOW
NORMAL HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

FINAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

PREDICTED DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

(98.8%)

OTHER POSSIBLE DIAGNOSES:

NONE

PATIENT: 8

CLINICIAN: DR. GLANZ

DATE: 12/10/75

SYMPTOMS:

GE 37 WEEKS GESTATION
LE 3 DAYS OLD
MALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
RESPIRATORY RATE LE 70
CYANOSIS
PO2 INCREASE WITH 100% O2 GE 30
PH GE 7.30
S2 SINGLE
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO DIASTOLIC MURMER
HOLOSYSTOLIC MURMER AT THE LEFT LOWER STERNAL BORDER
EJECTION MURMER AT THE LEFT UPPER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL GT 0.06
AXIS 181-269
P/S IN V1 GT 10
INCREASED PULMONARY BLOOD FLOW WITH VENOUS OBSTRUCTION
MODERATELY INCREASED HEART SIZE

CLINICAL DIAGNOSIS:

TRANSPOSITION OF THE GREAT ARTERIES

FINAL DIAGNOSIS:

TRANSPOSITION OF THE GREAT ARTERIES

PREDICTED DIAGNOSIS:

TRANSPCISITION OF THE GREAT ARTERIES (86.6%)

OTHER POSSIBLE DIAGNOSES:

VENTRICULAR SEPTAL DEFECT (5.8%)
HYPOPLASTIC LEFT HEART (4.0%)
TRUNCUS ARTERIOSUS (1.9%)

PATIENT: 9

CLINICIAN: DR. RESIDENT

DATE: 11/17/75

SYMPTOMS:

GE 37 WEEKS GESTATION
LE 3 DAYS OLD
MALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE GT 70
NO CYANOSIS
PH GE 7.30
S2 SPLIT
S3 PRESENT
S4 ABSENT
EJECTION CLICK ABSENT
NO DIASTOLIC MURMER
SYSTOLIC MURMER AT THE LEFT LOWER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
T IN V1 GT 0 IN PATIENT AGE GT 3 DAYS
NORMAL PULMONARY BLOOD FLOW
MODERATELY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

COARCTATION WITH A VENT SEPTAL DEFECT

FINAL DIAGNOSIS:

COARCTATION WITH A VENT SEPTAL DEFECT

PREDICTED DIAGNOSIS:

COARCTATION WITH A VENT SEPTAL DEFECT (49.1%)

OTHER POSSIBLE DIAGNOSES:

COARCTATION OF THE AORTA (24.3%)
VENTRICULAR SEPTAL DEFECT (15.7%)
AORTIC STENOSIS (5.0%)
PATENT DUCTUS ARTERIOSUS (4.1%)

PATIENT: 10

CLINICIAN: DR. RESIDENT

DATE: 9/30/75

SYMPTOMS:

LE 36 WEEKS GESTATION
GE 4 DAYS OLD
FEMALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
RESPIRATORY RATE GT 70
NO CYANOSIS
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
CONTINUOUS MURMUR AT THE APEX
CONTINUOUS MURMUR AT THE LEFT LOWER STERNAL BORDER
CONTINUOUS MURMUR AT THE LEFT UPPER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
P/S IN V1 LT 0.8
T IN V1 GT 0 IN PATIENT AGE GT 3 DAYS
INCREASED PULMONARY BLOOD FLOW W/O VENOUS OBSTRUCTION
NORMAL HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

FINAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

PREDICTED DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

(97.0%)

OTHER POSSIBLE DIAGNOSES:

VENTRICULAR SEPTAL DEFECT

(1.5%)

PATIENT: 11
CLINICIAN: DR. NUDEL
DATE: 12/28/75
SYMPTOMS:

GE 37 WEEKS GESTATION
LE 3 DAYS OLD
FEMALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELCW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE LE 70
CYANOSIS
PO2 INCREASE WITH 100% O2 LT 30
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO SYSTOLIC MURMER
NO DIASTOLIC MURMER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 181-269
R/S IN V1 LE 10 AND GE 0.8
NORMAL PULMONARY BLOOD FLOW
MODERATELY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

NON-CARDIAC DISEASE

FINAL DIAGNOSIS:

NON-CARDIAC DISEASE

PREDICTED DIAGNOSIS:

TRANSPOSITION OF THE GREAT ARTERIES (35.1%)

OTHER POSSIBLE DIAGNOSES:

NON-CARDIAC DISEASE (30.7%)
PERSISTENT FETAL PATHWAYS (16.9%)
TOTAL ANOMOLOUS PULM VENOUS DRAINAGE (11.5%)
PATENT DUCTUS ARTERIOSUS (1.6%)
VENTRICULAR SEPTAL DEFECT (1.1%)

PATIENT: 12

CLINICIAN: DR.

DATE: 00/00/00

SYMPTOMS:

LE 36 WEEKS GESTATION
GE 4 DAYS CLD
FEMALE
PULSE RATE LE 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LF 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE LE 70
CYANOSIS
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO DIASTOLIC MURMER
HOLOSYSTOLIC MURMER AT THE LEFT LOWER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 6C-180
P/S IN V1 LE 10 AND GE 0.8
T IN V1 LE 0 IN PATIENT AGE GT 3 DAYS
INCREASED PULMONARY BLOOD FLOW W/O VENOUS CBSTRUCTION
MODERATELY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

VENTRICULAR SEPTAL DEFECT

FINAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

PREDICTED DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

(49.7%)

OTHER POSSIBLE DIAGNOSES:

VENTRICULAR SEPTAL DEFECT

(49.4%)

APPENDIX 5 : Sample Patients-Reanalyzed with
Incidence Data Omitted



PATIENT: 2
CLINICIAN: DR. GLANZ
DATE: 10/26/75

SYMPTOMS:

GE 37 WEEKS GESTATION
GE 4 DAYS OLD
FEMALE
PULSE RATE GT 170
NORMAL PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER GT 2 CM BELOW CM
RESPIRATORY RATE GT 70
NO CYANOSIS
PH GE 7.30
S2 SPLIT
S3 ABSENT
S4 ABSENT
EJECTION CLICK PRESENT
NO DIASTOLIC MURMER
SYSTOLIC MURMER AT THE LEFT LOWER STERNAL BORDER
SYSTOLIC MURMER AT THE LEFT UPPER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 181-269
P/S IN V1 LE 10 AND GE 0.8
T IN V1 GT 0 IN PATIENT AGE GT 3 DAYS
INCREASED PULMONARY BLOOD FLOW W/O VENOUS CBSTRUCTION
MODERATELY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT

FINAL DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT

PREDICTED DIAGNOSIS:

ENDOCARDIAL CUSHION DEFECT (29.6%)

OTHER POSSIBLE DIAGNOSES:

VENTRICULAR SEPTAL DEFECT (20.4%)
TOTAL ANOMALOUS PULM VENOUS DRAINAGE (14.9%)
COARCTATION WITH A VENT SEPTAL DEFECT (11.2%)
AORTIC STENOSIS (7.7%)
PULMONARY STENOSIS (6.4%)
TRICUSPID ATRESIA WITH VSD (3.3%)
PATENT DUCTUS ARTERIOSUS

PATIENT: 5
CLINICIAN: DR. TALNER
DATE: 11/25/75
SYMPTOMS:

GE 37 WEEKS GESTATION
GE 4 DAYS OLD
MALE
PULSE RATE LE 170
INCREASED PULSE INTENSITY
NO DIFFERENTIAL PULSES
LIVER LE 2 CM BELOW CM
NO DIFFERENTIAL CYANOSIS
RESPIRATORY RATE GT 70
CYANOSIS
PO2 INCREASE WITH 100% O2 GE 30
PH GE 7.30
S2 SINGLE
S3 ABSENT
S4 ABSENT
EJECTION CLICK ABSENT
NO DIASTOLIC MURMUR
EJECTION MURMUR AT THE LEFT UPPER STERNAL BORDER
PR INTERVAL LE 0.11
QRS INTERVAL LE 0.06
AXIS 60-180
R/S IN V1 LE 10 AND GE 0.8
T IN V1 LE 0 IN PATIENT AGE GT 3 DAYS
INCREASED PULMONARY BLOOD FLOW WITH VENOUS OBSTRUCTION
MARKEDLY INCREASED HEART SIZE
LEFT AORTIC ARCH

CLINICAL DIAGNOSIS:

PATENT DUCTUS ARTERIOSUS

FINAL DIAGNOSIS:

TRUNCUS ARTERIOSUS

PREDICTED DIAGNOSIS:

TRUNCUS ARTERIOSUS

(77.8%)

OTHER POSSIBLE DIAGNOSES:

VENTRICULAR SEPTAL DEFECT	(7.5%)
PATENT DUCTUS ARTERIOSUS	(5.5%)
COARCTATION OF THE AORTA	(3.4%)
TRANSPOSITION OF THE GREAT ARTERIES	(1.7%)
COARCTATION WITH A VENT SEPTAL DEFECT	(1.1%)

YALE MEDICAL LIBRARY

Manuscript Theses

Unpublished theses submitted for the Master's and Doctor's degrees and deposited in the Yale Medical Library are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but passages must not be copied without permission of the authors, and without proper credit being given in subsequent written or published work.

This thesis by _____ has been used by the following persons, whose signatures attest their acceptance of the above restrictions.

NAME AND ADDRESS

DATE

