1973

A Survey of Medical Diagnostic Software

Steven Stein
University of Central Florida

A SURVEY OF

MEDICAL DIAGNOSTIC SOFTWARE

BY

STEVEN STEIN
B.S., Florida Technological University, 1970

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in
Engineering in the Graduate Studies
Program of
Florida Technological University, 1973

Orlando, Florida

TABLE OF CONTENTS

INTRODUCTION 1

PART I. METHODS USED IN MEDICAL DIAGNOSTIC SOFTWARE

Chapter

I. DECISION TABLES: USE IN MEDICAL DIAGNOSTIC SOFTWARE . . 9

II. STATISTICAL THEORY IN MEDICAL DIAGNOSTIC SOFTWARE 14

1. Bayesian Probability
2. Discriminant Analysis

III. SYMPTOM-DISEASE COMPLEX: USE IN MEDICAL DIAGNOSTIC SOFTWARE 18

PART II. A LITERATURE CROSS-SECTION OF MEDICAL DIAGNOSTIC SOFTWARE

IV. MULTIORGAN DIAGNOSTIC SOFTWARE 24

1. Multiphasic Health Screening
2. Medical Data Screening

V. SINGLEORGAN DIAGNOSTIC SOFTWARE 36

1. Computer Diagnosis in Heart Disease
   A. Tulane University Study
   B. University of Missouri Study
2. Computer Diagnosis in Chromosome Analysis
3. Computer Diagnosis in Differential Preoperative Diagnosis for Pelvic Surgery
4. Computer Diagnosis in Evaluation of Electrolyte and Base Disorders
5. Computer Diagnosis in Thyroid Disease
   A. University of Florida Studies
   B. University of Texas Medical School Studies
   C. Cambridge Hospital Studies
   D. University of Arkansas Medical Center Study
6. Computer Diagnosis in Dermatology
7. Computer Diagnosis in Primary Bone Tumors

PART III. SURVEY ON MEDICAL DIAGNOSTIC SOFTWARE

VI. SURVEY OF THE ORLANDO, FLORIDA PHYSICIANS CONCERNING USE OF MEDICAL DIAGNOSTIC SOFTWARE 90

DISCUSSION OF RESULTS AND CONCLUSIONS 93

APPENDIX 98

BIBLIOGRAPHY 101

A SURVEY ON MEDICAL DIAGNOSTIC SOFTWARE

Mr. Steven Stein

ABSTRACT

The field of medical diagnostic software is reviewed to define its status in the medical profession. This is accomplished in a two-step procedure. The first step is a cross-section of literature on the topic, and the second step is a survey of physicians in a sample area.

The cross-section of literature presents some of the more advanced studies which have been conducted on medical diagnostic software. Also presented is an explanation of the logic used in diagnostic software and the results of several test cases.

The survey was of physicians in the Orlando, Florida area to define the actual application of medical diagnostic software. It presented a sample of physicians' feeling concerning the present use of medical diagnostic software. From these steps, the present status of medical diagnostic software is defined and projections concerning its future made.

Approved by _____

Director of Thesis

INTRODUCTION

Medical diagnostic software is a term which refers to computer programs that accept signs and symptoms as input and offer the diagnosis of a probable disease(s) as output. The basic idea of medical diagnostic software is to use data obtained from studies of previously encountered patients to predict a most likely disease for new patients. The relationship which exists between the signs, symptoms, diagnostic software and probable disease is shown in the block diagram in Figure 1.

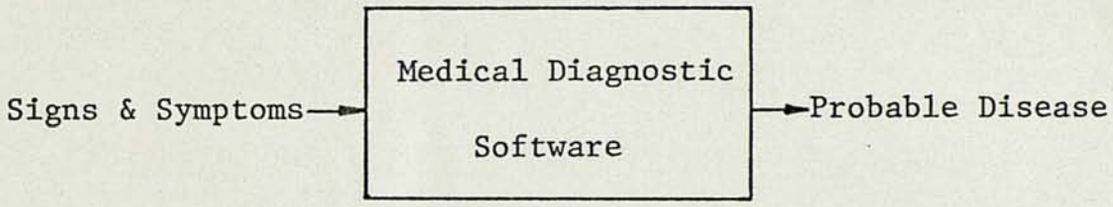

Figure 1

The signs and symptoms shown as input to the software (box) are recent clinical and historical information concerning the patient being considered. Medical diagnostic software (represented by the box) consists of mathematical equations and stored data that correlate signs and symptoms with probable diseases. The numerical values (stored data) used in the equations are calculated by using information collected from previously encountered patients in which trends could be established to correlate certain signs and symptoms with a disease. The disease diagnosed is not considered a definite diagnosis but instead a probable diagnosis based on data collected and stored as

part of the program from patients with similar symptoms and known treated diseases.

Medical diagnostic software is a tool which will allow a standard logical analysis to diagnose disease. The possibility exists that two physicians could diagnose two different diseases, both having the same clinical and historical information. Instead of a physician diagnosing a disease using his own experience and knowledge, diagnosis could be made by pooling together the experience and knowledge of many physicians stored in the memory of a computer which would be able to diagnose the same disease each time the symptoms of that particular disease were evident.

The lack of complete medical information about a patient's signs and symptoms as it relates to disease, much of which could be eliminated by use of a standard logical analysis, has been a major cause of misdiagnosis. The use of a computer to perform a large variety of predesignated arithmetical and logical instructions to make the correlation between signs, symptoms, and disease, provides physicians with a very powerful diagnostic aid. A brief review of the development of medical diagnostic software will assist in presenting its use as a diagnostic aid.

Medical diagnostic software as stated by Lee B. Lusted was an outgrowth of a theory that logical analysis for medical diagnosis was necessary, due to diagnostic errors which were caused by omissions of important medical data. Although the number of cases involving omissions was small, their importance when involved in human life must be considered.

One of the earliest aids which can be related to physicians striving for a logical analysis was a medical handbook. However, physicians would not use a handbook in the presence of their patients and, therefore, failed to use it while conducting examinations. Although this handbook was not used during diagnosis, it helped to make physicians realize that some type of device was needed to aid in a standard logical analysis.

F. N. Nash, an English physician, constructed a mechanical device early in the 1950's called a Logoscope, similar to a slide rule in structure, that allowed physicians to match combinations of eighty-two signs and symptoms, thus allowing the diagnosis of some 337 possible diseases.[1] Due to the limited number of possible diseases on Nash's Logoscope, investigation of a better method continued, and headed in another direction.

During the middle and late 1950's,[2] Paycha, Lipkin, and Hardy tried to develop another mechanical method using a card file method. Investigation continued throughout the 1950's, and a new idea in logical analysis, the use of automatic data processing in diagnosis, was developed.

Research on a logical analysis system for medical diagnosis assisted by automatic data processing began in 1957 by Ledley and

---

[1]Lee B. Lusted, "Computer Techniques in Medical Diagnosis," in Computers in Biomedical Research, ed. by Ralph W. Stacey, and Bruce D. Waxman. I (New York: Academic Press, 1965), 319.

[2]Ibid.

Lusted. They designed a system which was one of the first attempts to relate symptom complexes and disease complexes by conditional probability (Bayes' Theorem). They made an unsuccessful attempt to apply their theory on a group of lung cancer cases from the University of Rochester School of Medicine. A few years later, in 1961, Warner showed that Bayes' Theorem of conditional probability could be used successfully in the diagnosis of congenital heart disease.[3]

During the next several years, many new studies were conducted to relate computer systems to medical diagnosis. In fact, the beginning of the 1960's could be considered the "Industrial Revolution" of medical diagnostic software. By 1963 Lodwick had applied computer diagnosis to bone tumors; Rinaldo had applied computer diagnosis to epigastric pain; Overall and Gorhan had applied computer diagnosis to classification of psychiatric patients; Overall and Williams had applied computer diagnosis to thyroid disease; and Amosov, Shkabara, and Bykovskiy had applied computer diagnosis to heart disease.[4]

Researchers have been perfecting the techniques used in computer diagnosis during the last decade. In fact, some applications of computer diagnostic software are presently in use in large cities (Multiphase Health Testing), research hospitals, and teaching hospitals. The future of computer diagnostic software seems very promising; but before one could begin to speculate about its future, its present status in the medical profession should be evaluated by

---

[3] Ibid.

[4] Ibid., p. 320.

reviewing some of the literature which has been written covering this topic.

This paper has been written to define the present role of computer diagnostic software in the medical profession, which will allow conclusions to be drawn concerning its future use. This is to be accomplished in a two-step procedure. The first step is to present a cross-section of literature on the topic, and the second step is to conduct a survey of physicians in a sample area.

The cross-section of literature section presents some of the more advanced studies which have been conducted on medical diagnostic software. An explanation of the logic used in diagnostic software and the results of several test cases are presented.

The second step is to conduct a survey of physicians in the Orlando, Florida, area to define the real-life application of medical diagnostic software. The survey will present a sample of physicians' feeling concerning the present use of medical diagnostic software. By using the knowledge gained from these two steps, the cross-section of literature and the survey, the present status of medical diagnostic software will be able to be defined and projections concerning its future should be able to be made.

It should become apparent that this method of diagnosis has not completely evolved. Although many years have been spent trying to develop medical diagnostic software, its age of maturity has not been reached. It has followed the "branching-out" growth within the medical profession, specializing in the different organs of the body, but the

sophistication and reliability aspects which have always been an intricate part of medicine seems to be lacking.

The adolescent state of medical diagnostic software can be shown to exist by considering its status in the free enterprise system. In our capitalistic enterprise, as a new product emerges from the development stage to a basic level of maturity, it is marketed so that all of its potential users become familiar with the existence of the product. However, this does not seem to be evident in the field of medical diagnostic software. The survey which has been conducted will show that there still are physicians who are not aware of its use. If there was confidence in the sophistication and reliability of computer diagnostic software, the product would be marketed and physicians would be familiar with its use and would have been exposed to its possible applications and benefits.

This lack of maturity of computer diagnostic software could have several sources. One possibility is that this topic has not been given sufficient time and resources to develop. Another possibility is that although sufficient time and resources have been given, its mere presence in the medical profession has not been accepted.

Although these thoughts should remain with the reader, it should become apparent that medical diagnostic software has not been given sufficient time and resources to develop, but that there is a distinct desire to see this method of diagnosis grow. The basic approach expressed by the authors of the literature, and the results of the survey, should amplify the concept of future growth which will be discussed in more detail in the "Discussion of Results and

The survey reflected a definite interest in medical diagnostic software with 61.5% of the physicians stating that they would be willing to use a computer to assist in diagnosis if it were economically feasible. Not only are they willing to use this aid, 48.7% of the physicians felt their patients would respond favorably to its use with only 28.2% expressing a negative acceptance and 23.1% unsure. With as little background information the physicians have actually received, the interest in this topic is considered good and implies success for the future.

The author of the future programs which are to be written should be aware of some of the answers received on the survey from the potential users of medical diagnostic software. 35.9% of the physicians felt that greater accessibility was necessary and 17.9% felt that better accuracy should be established. 71.8% of the physicians felt that these programs should aim their use at physicians, 17.9% felt it should be aimed at nurses, and 5.1% felt it should be aimed at technicians. Other comments can be reviewed in the appendix of this paper.

Definite conclusions about the future use of medical diagnostic software did not seem to be presented in the literature which was reviewed. There was, however, a positive need expressed for continued research. The survey, on the other hand, showed a definite desire on the part of the physicians to make use of this tool. The potential users seem very interested in this aid which should act as the force to see that this concept of diagnosis will develop.

The interest shown by the physicians will become stronger in the next few years as more of them become aware of the topic which will

in turn stimulate a greater amount of research. The old theory of supply and demand can be easily applied in medical diagnostic software, for if a great demand is established by physicians, concentrated studies would appear in an attempt to supply this demand.

Medical diagnostic software has come a long way in the past decade, but its potential growth over the next decade should prove to be one of the great milestones in the medical profession.

APPENDIX A

SURVEY AND COVER LETTER

May 6, 1972

Dear Sir:

I am a graduate student at Florida Technological University and am working on my thesis, which is to be entitled "A Survey of Medical Diagnostic Software."

One portion of my work involves a survey of central Florida physicians concerning the use of a computer to aid physicians in the diagnosis of diseases. Although this field is relatively new, a number of computer programs are presently available and have been found to have application in cases such as the diagnosis of dermatological disorders, etc.

My interest in the material contained in this survey is three-fold: first, to determine the extent of present use of such programs in the central Florida area; second, to find how existing programs might be improved to better serve the needs of central Florida physicians; and third, to find the necessary characteristics which future programs should have in order that they might be useful and acceptable to the medical profession.

Your assistance by completing the enclosed survey will be greatly appreciated. It has been designed to take only a few minutes and, if you wish, a copy of the results will be sent to you.

Thank you,

Steven Stein

COMPUTER ASSISTED DIAGNOSTIC SURVEY
(CENTRAL FLORIDA)

It has been widely publicized that because of the current physician to patient ratio, there is a need to assist the physician in making the most efficient use of his time that is possible. It is with this concern that I am aking you to answer the following questions:

1. ☐ Yes ☐ No Have you ever used a computer diagnostic program to assist you in diagnosis?

2. ☐ Yes ☐ No If not, do you know of a diagnostic computer program presently being used in this area?

3. ☐ Yes ☐ No Would you be willing to use a computer to assist you in diagnosis if it were economically feasible?

4 How could diagnostic programs presently being used be revised to be of more use to you? (Check)

wider application

better accuracy

greater accessibility

simpler to use

others _____

_____

5. _____ (A, B, C or others) Do you feel that these programs should aim their use at (A) physicians, (B) nurses, or (C) technicians.

6. What reaction do you feel your patients would have to your use of a computer to assist in diagnosis?

_____

_____

7. In general, what advice would you offer the author of a diagnostic program to help make his program a useful aid to physicians? _____

_____

8. ☐ Yes ☐ No Would you like a copy of the results of this survey?

BIBLIOGRAPHY

Becker, H. C.; Nettleton, W. J., Jr.; Meyers, P. H.; Sweeney, J. W.; and Nice, C. M., Jr. "Digital Computer Determination of a Medical Diagnostic Index Directly from Chest X-Ray Images." IEEE Transactions on Biomedical Engineering, XI (July, 1964), 67-72.

Bleich, Howard L. "The Computer as a Consultant." Seminars in Medicine of the Beth Israel Hospital (Boston), CCLXXXIV (January 21, 1971), 141-6.

Bleich, Howard L. "Computer Evaluation of Acid-Base Disorders." Journal of Clinical Investigation, XLVIII (September, 1969), 1689-96.

Brandt, Edward N.; Hagans, James A.; and Schottsteadt, William W. "The Computer as a Diagnostic Aid in Medicine." Journal of the Oklahoma State Medical Association, LV (May, 1962), 213-17.

Brodman, Keeve, and Goldstein, Leo A. "The Medical Data Screen." Archives of Environmental Health, XIV (June, 1967), 821-26.

Brodman, Keeve, and van Woerkon, Adrianus J. "Computer-Aided Diagnostic Screening for 100 Common Diseases." Journal of the American Medical Association, CXCVII (September 12, 1960), 901-5.

Collen, Morris F. "Multiphasic Screening as a Diagnostic Method in Preventive Medicine." Methods of Information in Medicine, IV (November 2, 1965), 70-4.

"The Computer as Doctor." Newsweek, December 27, 1971, pp. 52-3.

Fitzgerald, Lawrence T.; Overall, John E.; and Williams, Clyde M. "A Computer Program for Diagnosis of Thyroid Disease." The American Journal of Roentgenology Radium Therapy and Nuclear Medicine, XCVII (August, 1966), 901-5.

"The Gainesville Community Multiphasic Health Screening Center." n.d. (pamphlet)

Hammond, Allen L. "Image Analysis: Application to Automated Medical Diagnosis." Science, December 3, 1971, pp. 1011-12.

Ledley, Robert S. "Computer Aids to Medical Diagnosis." Journal of the American Medical Association, CXCVI (June 13, 1966), 933-43.

Ledley, Robert S. Use of Computers in Biology and Medicine. New York: McGraw-Hill Book Company, Inc., 1965.

Lincoln, Thomas L., and Parker, Roger D. "Medical Diagnosis Using Bayes Theorem." Health Services Research, (Spring, 1967), 34-45.

Lodwick, G. S.; Cosmo, L. M.; Smith, W. E.; Keller, R. F.; and Robertson, E. D. "Computer Diagnosis of Primary Bone Tumors: A Preliminary Report." Radiology, LXXX (February, 1963), 273-75.

Lusted, Lee B. "Computer Techniques in Medical Diagnosis." Computers in Biomedical Research. Edited by Ralph W. Stacey, and Bruce D. Waxman. Vol. I, New York: Academic Press, 1965.

Lusted, Lee B., and Ledley, Robert S. "Mathematical Models in Medical Diagnosis." Journal of Medical Education, XXXV (March, 1960), 214-22.

Neurath, Peter W.; Enslein, Kurt; and Mitchell, George W., Jr. "Design of a Computer System to Assist in Differential Preoperative Diagnosis for Pelvic Surgery." The New England Journal of Medicine, CCLXXX (April 3, 1969), 745-49.

Norins, Arthur L. "Computers in Dermatology." Archives of Dermatology, XC (November, 1964), 506-11.

Oddie, T. H. "Computer Diagnosis for Tests of Thyroid Function." Journal of Clinical Endocrinology and Metabolism, XXXII February, 1971), 167-72.

Overall, John E., and Williams, Clyde M. "Conditional Probability Programs for Diagnosis of Thyroid Functions." Journal of the American Medical Association, CLXXXIII (February 2, 1963), 307-13.

Reichertz, P. L., Winkler, C. and Kloss, G. "Experiments with a Computer Program for the Diagnosis of Thyroid Diseases." Conference on the Use of Computers in Radiology, Chicago, October, 1966, E62-69.

Robin, Noel L.; Refetoff, Samuel; and Selenkow, Herbert A. "The Computer in the Diagnosis of Thyroid Disease." Journal of Nuclear Medicine, XII (September, 1971), 620-24.

Ryan, G. A. and Monroe, K. E. Computer Assisted Medical Practice: the AMA's Role. Chicago: American Medical Association, Center for Health Services Research and Development, 1971.

Templeton, A. W.; Bryan, K.; Waid, R.; Townes, J.; Hoque, M.; and Dwyer, S. J. "Computer Diagnosis and Discriminate Analysis Decision Schemes." Radiology, XCV (April, 1970), 47-55.

Wartak, Josef. "A Practical Approach to Automated Diagnosis." IEEE Transactions on Bio-Medical Engineering, BME-XVII (January, 1970), 37-43.

Williams, Clyde M.; Fitzgerald, Lawrence T.; and Overall, John E. Recent Advances in Diagnosis of Cancer. Chicago: Year Book Medical Publishers, 1966.

Conclusions" portions of this thesis. A brief breakdown of the sections of this paper will assist the understanding by establishing a sequence to follow.

The thesis has been divided into four main sections. The first section presents several of the mathematical methods which are used frequently in medical diagnostic software. The second section presents a cross-section of literature on medical diagnostic software. The third section presents the results of a survey sent out to physicians in the Orlando area to determine the present status of computer diagnostic software in a sample area, and the fourth section is conclusions which the author has made concerning the role of computer diagnostic software in the medical profession.

PART I

METHODS USED IN MEDICAL DIAGNOSTIC SOFTWARE

There are several logic techniques which have been used extensively in the field of medical diagnostic software such as decision tables, statistical theory, and symptom-disease complex. A brief explanation of these three techniques is presented in this section to facilitate the understanding of their use in specific diagnostic software applications found in other sections of this paper.

CHAPTER I

DECISION TABLES: USE IN MEDICAL DIAGNOSTIC SOFTWARE

Josef Wartak, in his article "A Practical Application to Automated Diagnosis," presented a method for computer aided diagnosis using decision tables which will be reviewed in this section. Wartak feels that the most important advantage of this technique is the power it permits to express a complex problem of medical diagnosis in a simplified form.

The simplified form is called a "Decision Table" which is a table of decision-making processes, partitioned into a set of four interconnected tables. An example of a decision table is shown in Figure 2.[5]

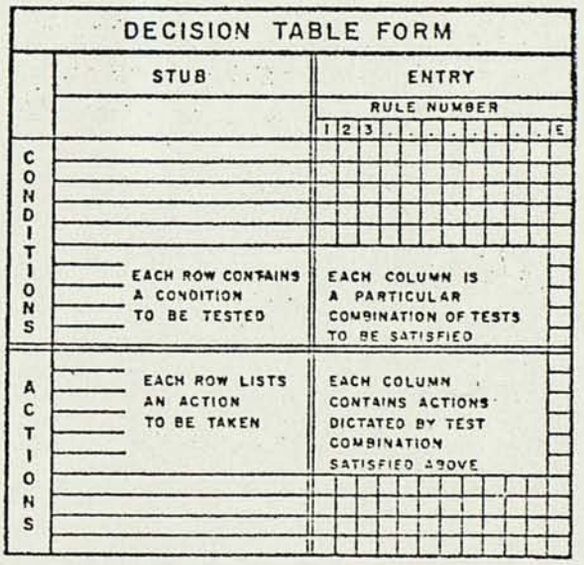

Figure 2. The skeleton outline of a decision table.

---

[5]Josef Wartak, "A Practical Approach to Automated Diagnosis," IEEE Transactions in Bio-medical Engineering, BME-XVII (January, 1970), 38.

The top left half of the table is conditions (symptoms) and the bottom left half is actions (cures or recommendations of additional tests). For example, the top left contains conditions to be tested such as sore throat, or fever, and the bottom left half contains actions to be taken as a result of the tests such as take two aspirins every four hours.

The table is completed by entering one of the decisions listed below which defines the conditions to be tested concerning the individual "rules:"

"Y" means Yes - test the condition to see if it is true;

"N" means No - test the condition to see if it is false;

" " means blank - this condition does not apply or this action is not to be taken when the rule is satisfied;[6]

"X" means X - this action is to be taken when all conditions for this rule are satisfied.

An example in the table below is given by considering "rule 1" which has three conditions to be tested to see if they are true: "QRS dur $\geq$ 0.12 sec in any 2 limb leads," "pr interval $\geq$ 0.12 in any 2 limb leads," "intrinsicoid deflection onset 0.06 sec in v1 or v2" (Figure 3).

The mechanics of the decision tables, as they might apply to medical diagnosis, can be shown by again considering Figure 3.[7] The

---

[6] Ibid.

[7] Ibid.

DECISION TABLE FOR DIAGNOSING ELECTROCARDIOGRAMS	TABLE NAME CONDUCTION DEFECTS	RULE NUMBER											
		1	2	3	4	5	6	7	8	E			
QRS DUR ≥ 0.12 SEC IN ANY 2 LIMB LEADS		Y	N	Y	N	Y	N	Y	N				
QRS DUR ≥ 0.10 SEC IN ANY 2 LIMB LEADS			Y		Y		Y		Y				
PR INTERVAL ≥ 0.12 SEC IN ANY 2 LIMB LEADS		Y	Y	Y	Y	Y	Y	N	N				
INTRINSICOID DEFLECTION ONSET ≥ 0.06 SEC IN ANY OF I, AVL, V5, V6		Y	N	N		N							
INTRINSICOID DEFLECTION ONSET ≥ 0.045 SEC IN ANY OF I, AVL, V5, V6			Y	N	N		N						
INTRINSICOID DEFLECTION ONSET ≥ 0.06 SEC IN V1 OR V2				Y	N	N							
INTRINSICOID DEFLECTION ONSET ≥ 0.05 SEC IN V1 OR V2					Y		N						
COMPLETE LEFT BUNDLE BRANCH BLOCK		X											
INCOMPLETE LEFT BUNDLE BRANCH BLOCK			X										
COMPLETE RIGHT BUNDLE BRANCH BLOCK				X									
INCOMPLETE RIGHT BUNDLE BRANCH BLOCK					X								
COMPLETE ATYPICAL BUNDLE BRANCH BLOCK						X							
INCOMPLETE ATYPICAL BUNDLE BRANCH BLOCK							X						
WOLFF-PARKINSON-WHITE SYNDROME								X	X				
GO TO TABLE INFARCTION PART I		X	X	X	X	X	X			X			
END OF CHECKING								X	X				

Figure 3. Decision table for conduction defects of the heart.

name of this table is "Conduction Defects" and refers to the diagnosis of electrocardiograms.

The first condition in the table, row one, is a check to see whether the QRS duration is greater than or equal to 0.12 seconds in any two of the limb leads. The outcomes for rules 1, 3, 5 and 7 are YES and for rules 2, 4, 6 and 8 are NO. If we consider rule 3 we see that YES answers were made for the first, third and sixth conditions while NO answers were made for the fourth and fifth conditions. Conditions two and seven do not apply for this rule. The actions for

this rule (rule #3) as can be seen in the bottom half of the table, are to "complete right bundle branch block" and "Go to Table Infarction Part I." Thus if we combine the conditions and actions associated to rule 3, it would read as follows:

> If QRS duration in any two limbs is greater than or equal to 0.12 second, and if PR interval in any two limb leads is greater than 0.12 second, and if intrinsicoid deflection onset in any lead of L1, AVL, V5, V6 is not greater or equal to 0.045 second, and if intrinisicoid deflection onset is V1 or V2 is greater than or equal to 0.06 second, then the diagnosis of complete right bundle branch block is issued and the checking of Infarction Part I Table is ordered.[8]

The actual application of decision tables in medical diagnosis can be shown by the following example. The clinical data for a patient reveals that the QRS duration is 0.12 seconds in the two limb leads, the PR interval is 0.12 seconds in two limb leads, and the intrinsicoid deflection onset was 0.02 seconds for V5 and 0.07 seconds for V2. Compare the clinical data to the conditions listed for each rule in the decision table starting with rule 1. If we compare the clinical data of the patient with rule 1 it is seen that condition three (condition on row 3) is not met; therefore go to rule 2. Check the clinical finding of our patient against the conditions of rule 2 and find that condition five is not met. This process is continued until a rule is found that matches the conditions of our patient. This is accomplished, in this example, when rule 3 is reached. Once we have found a rule that matches, the actions listed in the bottom half of the table are applied. Once a rule which matches the conditions of the particular patient is found, it is time to stop. If

---

[8] Ibid.

none of the rules are satisfied, the ELSE rule comes into action and the diagnostic process is transferred to "Table Infarction Part I."[9]

Numbers assigned to the rules do not imply that the rules must be executed in any given order. They are independent and may be selected in any sequence. Although the set of conditions for each rule is independent, they may lead to the same action.

The last rule in the right-hand corner is the ELSE rule which contains no entries in the condition, or upper part of the table. This rule specifies the action to be taken if no other rules are satisfied.

It might be convenient to think of this table as a matrix, for there is no limit to the number of rules and conditions and actions which can be used. However, the author points that for convenience it might be best to keep the table small and link the tables together by "GO TO" actions as is shown in the second to the last action of the decision table shown in Figure 3. There is no limit to the number of tables which could be linked together. Another option instead of "GO TO" is "DP." This would allow one table to execute another table and then return to the original table.

---

[9] Ibid.

CHAPTER II

STATISTICAL THEORY IN MEDICAL DIAGNOSTIC SOFTWARE

There are presently two main statistical tools being applied in the field of medical diagnostic software. They are the Bayesian Probability Method and the Discriminant Analysis Method.

The mathematical techniques of these two methods will be presented in this section with their application to individual software presented in the literature cross-section portion of this paper.

1. Bayesian Probability

Bayes' Theorem has been the most widely used statistical tool for computer aided medical diagnosis. Many of the studies presented in this paper use this technique.

The general formula for Bayes' Theorem as applied to the diagnosis of a disease with certain symptoms is given below. The probability that a patient who has a certain symptom has disease "r" is represented by

$$P(\text{Disease}_r \mid \text{Symptom}) = \frac{P(\text{Symptoms} \mid \text{Disease}_r) \times P(\text{Disease}_r)}{\sum_{i=1}^{N} [P(\text{Symptoms} \mid \text{Disease}_i) \times P(\text{Disease}_i)]}$$

where N is the number of possible diseases. If "S" denoted the set of symptoms and "D" denoted the set of diseases, then

$$P(D_r|S) = \frac{P(S|D_r) \times P(D_r)}{\sum_{i=1}^{N} [P(S|D_i) \times P(D_i)]}$$

where

N - is the number of possible diseases,

$P(D_i)$ - "is a prior probability of having 'disease i' for the population of which the patient is a member" (possibly a hospital's population), and

$P(S|D_i)$ - "is the probability of occurrence of a particular set of symptoms belonging to the patient, given a population with 'disease i.'"[10]

Brandt, Hagans, and Schottstaedt have pointed out two problems encountered with the Bayesian model. First, due to the scarcity of available data, there is difficulty in obtaining good probability estimates of the rarer diseases, and second, it is difficult to separate diseases (more than one) possessed by one patient. The probabilities are established for "disease i," that is, one disease with several symptoms.[11]

Examples of this technique are presented in the thyroid and bone tumor sections.

2. Discriminant Analysis

Brandt, Hagans and Schottsteadt in their article "The Computer as a Diagnostic Aid in Medicine" state that discriminant analysis is

---

[10]Thomas L. Lincoln and Rodger D. Parker, "Medical Diagnosis Using Bayes' Theorem," Health Services Research, (Spring, 1967), 34-5.

[11]Edward N. Brandt, James A. Hagans, and William W. Schottsteadt, "The Computer as a Diagnostic Aid in Medicine," Journal of the Oklahoma State Medical Association, LV (May, 1962), 216.

considered to approximate the clinical approach to disease diagnosis better than any other known model. The discriminant analysis technique, as applied to medical diagnosis, is accomplished by assigning patients individually to one discrete diagnostic category by considering their particular symptoms. The discrete diagnostic category, denoted by "Z" is calculated as a weighted sum of the clinical features of a patient's illness. Each patient is classified in a diagnostic category by adding up the weighted values associated with his symptoms. The general equation to determine the diagnostic category "Z" is

$$Z = B_i X_i$$

or for N symptoms

$$Z = B_i X_i = B_1 X_1 + B_2 X_2 + \ldots + B_N X_N$$

where the X's represent the symptoms and the B's are the weights attached to each symptom. The weights, $B_1 \ldots, B_N$, are determined by considering a sample of patients with known diseases. The weights are assigned in such a way that there will be maximum discrimination among the diagnostic categories. Therefore, the individual features which are highly associated with the presence of a disease will receive the largest weights. The symptoms (X) can be tested to determine the significance of its weight (B) in determining the diagnostic category (Z). Using the equation above, the clinical findings of a patient can be substituted in to determining the patient's diagnostic category,Z.

A graphical method is called Linear Decision Function (LDF) as presented by Templeton, Waid, Hoque, and Dwyer. The method "separates linear decision boundaries (hyperplanes) which separate pattern

classes (a boundary of line divides two classes; the lines dividing pattern classes form hyperplanes)." Positioning of the hyperplanes is accomplished by using samples in which the classification of disease is known.[12]

An example of this graphical method is presented in the computer diagnosis in differential preoperative diagnosis for pelvic surgery section of this paper.

---

[12]A. W. Templeton, et al., "Computer Diagnosis and Discriminate Analysis Decision Schemes," Radiology, XCV (April, 1970), 48.

CHAPTER III

SYMPTOM-DISEASE COMPLEX: USE IN MEDICAL DIAGNOSTIC SOFTWARE

Lusted and Ledley have presented numerous articles on the computational techniques of symptom-disease complex in medical diagnosis. Their model uses a binary notation (1,0) or (YES,NO) in a table format to analyze patients' data. This data can then be transferred to a digital computer.

A symptom-disease complex is a list of symptoms and a list of diseases which may possibly be correlated, one with the other, by considering all of their possible combinations.

All possible combinations of a symptom-disease complex, using a binary system of "1,0" are placed in a table. ("0" indicates that it does not occur and "1" that it does occur.) The number of possible combinations is 2^n where n is the total number of diseases plus the total number of symptoms. If we select two diseases, D(1), and D(2), and two symptoms, S(1), and S(2), we have 2^4 or 16 possible symptom-disease complexes as can be seen in Table I. The table is set up with the two symptoms and the two diseases -- S(1), S(2), D(1), D(2) -- listed as the rows with the 16 possible combinations of complexes as columns.

TABLE I

Logical Basis for S(1), D(1), and D(2)

	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4
S(1)	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
S(2)	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
D(1)	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
D(2)	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
			d_1				d_2				d_3				d_4	

Many of the combinations which have been listed in Table I do not occur and can be eliminated due to the conditions (medical history) known about the diseases and symptoms.[13] For instance, one of the conditions is that in order to have D(2), symptom S(1) must be present, or mathematically, $D(2) \rightarrow S(1)$. Each time a "1" appears in row D(2) there must be a "1" in row S(1). Otherwise that column is eliminated from the table. Therefore, columns S_1d_3, S_3d_3, S_1d_4, and S_3d_4 can be eliminated. We also have knowledge that if D(1) is present and D(2) is not, then S(2) must be present, or mathematically, $D(1) \cdot \overline{D(2)} \rightarrow S(2)$. Each time a "1" appears in row D(1) and a "0" in row D(2) there must be a one in row S(2). Therefore, the conditions of columns S_1d_2 and S_2d_2 are violated and can also be eliminated. With the condition $\overline{D(1)} \cdot D(2) \rightarrow S(2)$ columns S_3d_3 and S_4d_2 can be eliminated and columns S_2d_1, S_3d_1, and S_4d_1 are eliminated by the condition $S(1) + S(2) \rightarrow D(1) + D(2)$.

---

[13]Lee B. Lusted and Robert S. Ledley, "Mathematical Models in Medical Diagnosis," Journal of Medical Education, XXXV (March, 1960), 216.

Complexes can be eliminated until there remains a reduced table which contains all the combinations permitted with the medical knowledge available about the diseases and symptoms with which we are dealing. Table II shows "the reduced basis for S(1), S(2), D(1) and D(2) which includes medical knowledge."[14] The medical knowledge known established the rules which allowed certain columns to be eliminated.[15]

TABLE II

The Reduced Basis for S(1), S(2), and D(1), and D(2) Which Includes Medical Knowledge

	S_1	S_3	S_4	S_2	S_2	S_4
S(1)	0	0	1	1	1	1
S(2)	0	1	1	0	0	1
D(1)	0	1	1	0	1	1
D(2)	0	0	0	1	1	1
	d_1	d_2		d_3	d_4	

If we now analyze (using Table II) a patient with symptoms S(1), S(2), a "1" in rows S(1) and S(2), it is only possible for him to have S_4d_2 or S_4d_4. All other columns can be eliminated. If he has S_4d_2, then he possesses D(1), and if he has S_4d_4, he possessed D(1) and D(2). Therefore, we can conclude that the patient has D(1) but no other conclusions can be drawn until additional knowledge is added to the table or more tests are performed on the patient.

---

[14] Ibid., p. 217.

[15] Ibid.

The authors point out that this technique has received "little attention in medical texts and in medical student teaching in this country."[16]

---

[16] Ibid., p. 221.

PART II

A LITERATURE CROSS-SECTION OF MEDICAL DIAGNOSTIC SOFTWARE

The topic of medical diagnostic software covers a wide variety of different programs which have been used to diagnose diseases. These programs have been directed at correlating symptoms to disease by using mathematical models.

The literature which has been written on medical diagnostic software reflects two main divisions or areas of interest.

The first area of interest is titled Multiorgan Diagnostic Software and refers to that software which can diagnose disease in several organs or sections of the body. General medical information gathered for a general checkup is the input for multiorgan diagnostic software.

The second area is Singleorgan Diagnostic Software which is limited to diagnose disease in one particular organ or section of the body. When a particular organ or section is believed to be diseased, information concerning this section of the body would be the input for single organ diagnostic software. The singleorgan diagnostic software should not be visualized as a simplified version of multiorgan diagnostic software. Instead of having detailed information concerning many organs, singleorgan diagnostic software deals with a lot of information concerning one particular section.

With this separation in mind, this section has been divided into two areas, Multiorgan Diagnostic Software and Singleorgan Diagnostic Software.

CHAPTER IV

MULTIORGAN DIAGNOSTIC SOFTWARE

Multiorgan diagnostic software refers to computer programs that diagnose disease in several sections of the body. Two examples of this software titled Multiphasic Health Screening and Medical Data Screening are presented in this section.

1. Multiphasic Health Screening

Multiphasic health screening is a method for providing comprehensive, quantitative and precise testing for a large number of diseases. The first multiphasic screening center was developed under the Kaiser-Permanente Program in Oakland, California, in 1951.[17] This particular system, called Automatic Multiphasic Health Testing (AMHT), now receives over 30,000 adults each year.[18]

The American Medical Association considers the AMHT program an economical and effective method of patient management.[19] It permits physicians to see large numbers of patients and is viewed as a potential method for use in preventive medicine and for the early detection of disease. The complete Multiphasic Health checkup consists of three

---

[17]"The Gainesville Community Multiphasic Health Screening Center." n.d. (pamphlet)

[18]Morris F. Collen, "Multiphasic Screening as a Diagnostic Method in Preventive Medicine," Methods of Information in Medicine, IV (November 2, 1965), 71.

[19]G. A. Ryan and K. E. Monroe, Computer Assisted Medical Practice: the AMA's Role. (Chicago: American Medical Association, Center for Health Services Research and Development, 1971), p. 1.

parts: (1) the automated Multiphasic Screening Program (tests and procedures); (2) physical examination by an internist; and (3) a group of examinations by specialists (physicians).

The test portion, part 1, which involves the computer diagnosis and therefore the focus of attention in this discussion, takes two hours and consists of those phases listed below.

1. Electrocardiogram (6 leads) and phonocardiogram (at apex and base), simultaneously graphed on paper by a direct recording, multiple channel Sanborn optical oscillograph.
2. Table tilt cardiovascular test (80° tilt in 35 sec.; pulse and blood pressure recorded supine and 1 minute after tilt).
3. Height, weight, and other body measurements, recorded by an automated anthropometer with direct punched card output.
4. Chest x-ray (70 mm. minifilm).
5. Breast x-ray (cephalocaudal and lateral views in women).
6. Visual acuity (modified Sloan chart) and pupillary escape test (an unsustained pupillary light reflex occurring in retinal or optic nerve disease).
7. Ocular tension (measured by Schiotz tonometer).
8. Retinal photograph (of left eye).
9. Timed vital capacity (1 sec. and total) by the Gaensler-Collins spirometer.
10. Pain reaction test (a modified Libman test measured as pain tolerance to increasing pressure on Achilles tendon).
11. Hearing test (for 6 frequencies by a Rudmose-Bekesy automated audiometer).
12. Self-administered health questionnaire form (present and past history). In addition, a set of 207 medical questions on pre-punched cards are sorted by the patient so that the "yes" responses can be automatically repro0 duced for computer processing.
13. Personality appraisal questionnaire (a modified MMPI type of test with 207 psychological questions sorted by the patient in a manner identical to (12) above.
14. Tetanus toxoid immunization (with Hypospray gun).
15. Blood chemistries performed within 12 minutes by a multichannel Technicon AutoAnalyzer with direct punch card output, including:
    a) Serum glucose (1 hour after 100 gms. glucose).
    b) Serum creatinine.

c) Serum albumin.
d) Serum total protein.
e) Serum cholesterol.
f) Serum uric acid.
g) Serum calcium.
h) Serum transaminase (SGOT).

16. Blood hemoglobin and white cell count automatically measured.
17. Blood group (AB O)
18. Serological (V.D.R.L.) test for syphilis.
19. Blood rheumatoid factor (latex test).
20. Urinanalysis for:
    a) Urine pH, glucose, protein, and blood (paper strip) tests.
    b) Urine bacteria (4 hour triphenyltetrazolium chloride culture of midstream specimen).
21. In accordance with programmed "advice" rules, additional tests are arranged by means of an IBM 1440 computer, including return appointments for physicians. All test results are stored for summary report printout.[20]

Most of the data generated in the testing (those listed above) is recorded on pre-punched or marked cards. Other information from tests such as cardiograms and roentgenogiams is interpreted by the physician and must be entered manually. At other phases of the test technicians record information on marked-sensed cards which accompany the patient, together with a header card, throughout the testing. Another phase in testing is the questionnaire station where patients are given a letter box divided into three sections. The top section contains a deck of 207 computer cards pre-punched, each asking a single question. The patient responds by taking a card out of the top section and dropping it into the middle section if the answer to the

---

[20]Morris F. Collen, "Multiphasic Screening as a Diagnostic Method in Preventive Medicine," Methods of Information in Medicine, IV (November 2, 1965), 71-2.

question is no. The laboratory phase consists of automatic punched cards with the chemistry tests and urine tests sent by pneumatic tube to the computer input station.

At the last phase (station), the cards which the patient carried with him are forwarded to the data processing room where the sensed cards are punched and sent to the computer station.

This information obtained from the computer output may advise a receptionist to arrange for certain additional tests to be performed.

The reports for such tests as x-ray, mamogram and electrocardiogram are forwarded to the computer station after a one or two day lag, at which time all the data is placed in the computer. The output is a summary report of the screening test, including provisional diagnoses and listings of any additional tests and procedures necessary to make a more accurate diagnosis. A diagnosis, which may or may not require additional medical attention, can then be related to the patient.[21]

There are presently 122 active or planned multiphasic programs in operation and they are listed in the Multiphasic Screening and Automated Health Evaluation Program Directory. The average fee is about $33 per person but may range from $5 to $225. Dr. Morris F. Collen has stated that these systems are feasible and can be implemented within a community of physicians prepared to support the system.[22]

---

[21] Ibid., pp. 71-4.

[22] Ryan and Monroe, "Computer Assisted Medical Practice," p. 1-2.

2. Medical Data Screening

Brodman and Goldstein have presented a method to aid physicians in the diagnosis of 100 common diseases by use of a questionnaire to collect the patient's history and a subsequent computer analysis of reported symptoms. This method has proven 70 percent correct in identifying diseases in employees of an insurance firm, and has shown similar results among private patients of internists and a general practitioner. The MDS attempts to screen a large group of people and identify them with one of 100 common diseases listed below.[23]

Code	Disease
	Eye
A370	Conjunctivitis
A380	Refractive error
A384	Strabismus
A385	Cataract
A387	Glaucoma
	Ear
B390	Otitis externa
B391,2	Otitis media
B398	Deafness
	Buccal Cavity
C210	Benign buccal neoplasm
C530	Dental caries
C531,2	Disorder support of teeth
C533	Disorder of occlusion
	Respiratory System
D162,3	Malignant neoplasm of lung
D240	Hay fever
D241	Asthma
D470,1,2,3,4,5	Acute upper respiratory infection
D502	Chronic bronchitis
D510	Hypertrophy of tonsils
D512	Chronic pharyngitis
D513	Chronic sinusitis
D514	Deflected nasal septum
D515	Nasal polyp

---

[23]Keeve Brodman and Adrianus J. van Woerkom, "Computer Aided Diagnostic Screening for 100 Common Diseases," Journal of the American Medical Association, CXCVII (September 12, 1966), 901.

	Circulatory System
E410,1,2	Chronic rheumatic heart disease
E420	Arteriosclerotic heart disease
E433	Functional disease of heart
E440	Hypertension with heart disease
E444	Hypertension
E450	General arteriosclerosis
E453	Peripheral vascular disease
	Digestive System
F151	Malignant neoplasm of stomach
F153,4	Malignant neoplasm of large intentine
F540	Ulcer of stomach
F541	Ulcer of duodenum
F544	Disorder function of stomach
F572	Chronic enterisis
F573	Functional disorder of intestines
F574	Anal fissure
F581	Cirrhosis of liver
F584	Cholelithiasis
	Urinary System
G600	Infection of kidney
G602	Calculi of kidney
G605	Cystitis
	Male Genital Organs
H610	Hyperplasia of prostate
H611	Prostatitis
H613	Hydrocele
H614	Orchitis
	Female Genital Organs
I170	Malignant neoplasm of breast
I213	Benign neoplasm of breast
I214	Uterine fibromyoma
I215	Other benign neoplasm of uterus
I620	Chronic cystic disease of breast
I630	Infective disease of uterus
I631	Uterovaginal prolapse
I634	Disorders of menstruation
I635	Menopausal symptoms
	Bones and Organs of Movement
J722	Rheumatoid arthritis
J723	Osteo-arthritis
J735	Displacement of intervertebral disc
J741	Synovitis
J745	Curvature of spine
	Skin and Cellular Tissue
K131	Dermatophytosis
K190,1	Malignant melanoma
K220	Benign melanoma
K221	Pilonidal cyst
K222	Benign neoplasm of skin

K226	Lipoma
K228	Hemangioma
K690	Boil
K691	Cellulitis of finger or toe
K696	Infectious wart
K700	Seborrheic dermatitis
K701	Eczema
K705	Erythematous condition
K706	Psoriasis
K708	Pruritus
K713	Disease of hair
K714	Disease of sebaceous glands
	Nervous System
L353	Epilepsy
L354	Migraine
L791	Headache
	Miscellaneous Diseases
M020,8	Syphilis
M030	Gonorrhea
M250	Simple goiter
M251	Nodular goiter
M252	Thyrotoxicosis
M253	Myxedema
M260	Diabetes mellitus
M287	Obesity
M290,1,2,3	Anemia
M322	Use of alcohol
M460	Varicose veins
M461	Hemorrhoids
M463	Phlebitis
M560	Hernia
	Psychoneurotic Disorders
N310	Anxiety reaction
N311	Hysterical reaction
N314	Neurotic-depressive reaction
N315	Somatization reaction, circulatory
N316	Somatization reaction, digestive
N318	Psychoneurotic disorder mixed[24]

A patient's symptoms are reported by having the patient answer questions on the Medical Data Index-Health Questionnaire, which consists of 150 yes-no medical questions and takes about fifteen minutes to complete. The Male and Female titled question forms differ on eight

---

[24]Keeve Brodman and Leo A. Goldstein, "The Medical Data Screen," Archives of Environmental Health, XIV (June, 1967), 822-23.

questions concerning the genital system. A patient's symptoms are matched with other patients' symptoms of the same sex who were diagnosed by physicians as possessing a certain disease. A resultant list of possible diseases is then established. This list is presented to the patient's physician. The report only informs the physician that a patient has complexes of symptoms commonly found in the disease named. It is necessary for the physician to draw his own conclusions to determine why the patient claimed the particular symptoms. This might require the physician to perform an interrogation of the patient, a physical examination, obtain advice from consultants, or gather information concerning the patient's family, social or work history. Until a physician can confirm one of the diseases, the computer list of diseases presents presumptive and not proven diagnoses.

The study reported by Brodman and Goldstein involves 115 New York City employees of the Home Life Insurance Company (almost all of whom were clerical or supervisory personnel) who completed the Medical Data Screen (MDS) questionnaire. The medical records of these employees were made available to a medical staff consisting of two physicians.

In order to offer a comparison with the group discussed above, two other groups were considered. One group involved 208 office patients under the care of four internists at the New York Hospital-Cornell Medical Center, and the second group involved 44 office patients of a New York City general practitioner. The average age in these two groups was a little greater than that in the insurance company.

Each of the patients was examined by the physicians of their particular groups and then a record was made of the physician's findings. The results were compared with the disease complexes (list of diseases) identified by the Medical Data Screening. The correlations of the diagnosis of the physicians to that of the MDS were divided into the following categories (the alphanumeric symbols are used in the tables which follow):

(A) The physician diagnosed the same disease identified by the MDS;
(B) The physician diagnosed a disease in the same organ system and differential as the MDS;
(C) The physician had no information as to whether or not the disease identified by the MDS was present;
(D) The physician had information that neither the disease identified by the MDS nor any other clearly related disease was present;
(E) The disease was diagnosed by the physician but was one of the 100 diseases stored as part of the MDS;
(F) The disease was not diagnosed by the physician but was one of the 100 diseases stored as part of the MDS.

Table III gives a summary of the results by giving the average number of diseases per patient for all diseases identified by the MDS method (row 1), all diseases diagnosed by physicians (row 2), diseases both identified and diagnosed (row 3), and then a breakdown of the diseases only identified by the MDS technique (rows 4, 5 and 6) and only by the physicians (row 7 and 8), according to the three types of patients being cared for by the three types of physicians.[25] For example, an average of .1 diseases per patient was identified by the MDS technique and was known not to exist by the industrial physicians.

All diseases diagnosed by the physicians and the MDS method

---

[25] Ibid., p. 824.

were lower in the group of employees from the Home Life Insurance Company (Industrial) than the other two groups. This is probably due to the fact that this was the only group of the three that was not seeking medical care. The other two groups were selected from larger groups that were receiving medical opinion and/or treatment.

TABLE III

Average Numbers of Diseases Per Patient, by Type of Physicians Caring for Patient

	Symbol	Type of Physician Caring for Patient		
		Industrial Physician	General Practitioner	Internist
All diseases identified by MDS method	A+B+C+D	1.7	5.7	3.8
All diseases diagnosed by physician	A+E+F	1.1	2.4	2.9
Diseases both identified and diagnosed	A	0.8	1.7	1.7
Diseases only identified by MDS				
Differential to diagnosed diseases	B	0.2	1.5	1.1
Presence unknown to physician	C	0.7	2.4	0.4
Known by physician not to be present	D	0.1	0	0.5
Diseases only diagnosed by physician				
100 diseases of MDS	E	0.3	0.7	0.7
Other diseases	F	0.1	0	0.5
Number of cases		117	44	208

Ratios of the diseases per patient were derived from Table I and are shown in Table IV.[26] The mathematical explanations are shown in the symbol's column and use the alphanumeric symbols listed previously. In the category of the diseases diagnosed by the physicians, 70 percent, 71 percent and 59 percent identified by the MDS were diagnosed by the industrial physicians, general practitioner and internists, respectively.

The findings for individual patients (as opposed to the group findings) can be seen in Table V which is divided into three parts and

---

[26] Ibid.

shows cumulative percent distributions of diseases per patient for those jointly identified by MDS and diagnosed by physicians, identified by MDS, presence unknown to physicians and identified by MDS, known by physicians not to be present.[27]

TABLE IV

Ratios of Diseases Per Patient, by Type of Physician Caring for Patient

	Symbol	Type of Physician Caring for Patient		
		Industrial Physician	General Practitioner	Internist
Diseases diagnosed by physician				
% of all that are of 100	(A+E)/(A+E+F)	95	100	83
% of all identified by MDS	A/(A+E+F)	70.	71	59
% of 100 identified by MDS	A/(A+E)	73	71	71
Diseases identified by MDS				
% presence unknown to physician	C/(A+B+C+D)	39	42	11
% known not to be present	D/(A+B+C+D)	5	0	14
No. of cases		117	44	208

"In the first part of the table (Table V) only a few of the patients of the industrial physician had more than one diagnosis made for them." Part two of the table shows that the internists in .789 of the cases, the industrial physicians in .667 of the cases, and the general practitioners in .296 of the cases had knowledge of every disease identified by the screen. Part three shows that there were very few cases that the MDS identified a disease that the industrial physicians and the general practitioners did not know was present. This happened about one-fifth of the time with the internists.[28] (100% - 77.9% = 22.1% or 1/5)

---

[27] Ibid., p. 824.

[28] Ibid., p. 825.

TABLE V

Cumulative Percent Distribution, by Physician or Disease

Diseases Per Patient	Industrial Physicians	General Practitioners	Internists
Jointly identified by MDS and Diagnosed by Physician			
0	59	22.7	37
1	85.5	56.8	61.5
2	96	75	73.6
3	96.6	84.1	83.2
4	99.1	93.2	88.5
5	99.1	100	91.4
6	99.1	...	97.1
7 or more	100	...	100
No. of patients	117	44	208
Identified by MDS, Presence Unknown to Physician			
0	66.7	29.6	78.9
1	83.8	47.7	88.9
2	94.9	61.4	94.7
3	96.6	77.3	98.1
4	96.6	84.1	98.6
5	99.1	88.6	98.6
6	100	93.2	100
7 or more	...	100	...
No. of patients	117	44	208
Identified by MDS, Known by Physician not to be Present			
0	94	97.7	77.9
1	99.1	100	87.5
2	100	...	94.2
3	...	...	98.1
4	...	...	98.1
5	...	...	98.6
6	...	...	99.5
7 or more	...	...	100
No. of patients	117	44	208

CHAPTER V

SINGLEORGAN DIAGNOSTIC SOFTWARE

Singleorgan diagnostic software is a term used to describe applications of a computer program to diagnose disease in a single area of the body as opposed to multiorgan diagnostic software. Several of the more advanced areas of singleorgan diagnostic software are presented in this section.

1. Computer Diagnosis in Heart Disease

Application of the computer to diagnosis of heart disease has been assisted by the development of image analysis, the technique of converting an image to a two dimensional array of numbers representing the image and then analyzing it to obtain certain data. The studies to be discussed are experiments which have been taking place for several years at Tulane University and the University of Missouri.

A. Tulane University Study

The cardothoracic ratio, a standard test to measure enlargement of the heart, is the diagnostic criterion which was selected by a team at Tulane University.[29] The ratio is determined by dividing the maximum transverse diameter of the heart by the maximum transverse diameter of the thorax (the part of the body between the neck and

---

[29] H. C. Becker, et al., "Digital Computer Determination of a Medical Diagnostic Index Directly from Chest X-Ray Images," IEEE Transactions on Biomedical Engineering, XI (July, 1964), 67.

abdomen which contains the heart and lungs). This ratio is normally less than 50 percent, with an average of about 45 percent.

A photofluorogram, a photographic film image of a fluorescent screen widely used in mobile unit mass x-ray surveys, was used to produce an image matrix of thirty-seven adults which was then written on magnetic tape using an image-scanning system. The image scanning system separates the image information of each x-ray film into 160,000 discrete elements, represented by a two-digit number ranging from 00, full black, to 99, full white. The number represents 100 shades of gray which lie between complete darkness and complete transparency. "Although two-digit precision was used in the analog-to-digital conversion process, all 100 levels were not discretely reproducible because of the system-injected electronic noise and scanner fluorescent-screen nonuniformities." The image matrix of 160,000 discrete elements (which looks similar to an x-ray negative) can be thought of as a rectangular matrix with 320 rows and 500 columns. The rows represent 500 two-digit samples of the image x-ray and the columns are considered as a series of 320 digital records on magnetic tape. An example of a graph presented from the image matrix information is shown in Figure 4.[30]

The X addresses are the series of 320 digital records on magnetic tape, integrated at each X address with the sum shown under the curve with the intensities read on the Y axis. In other words, the

---

[30] Ibid., p. 69.

amplitude of the curve is the typical record showing integrated Y values at each of the 500 column addresses.

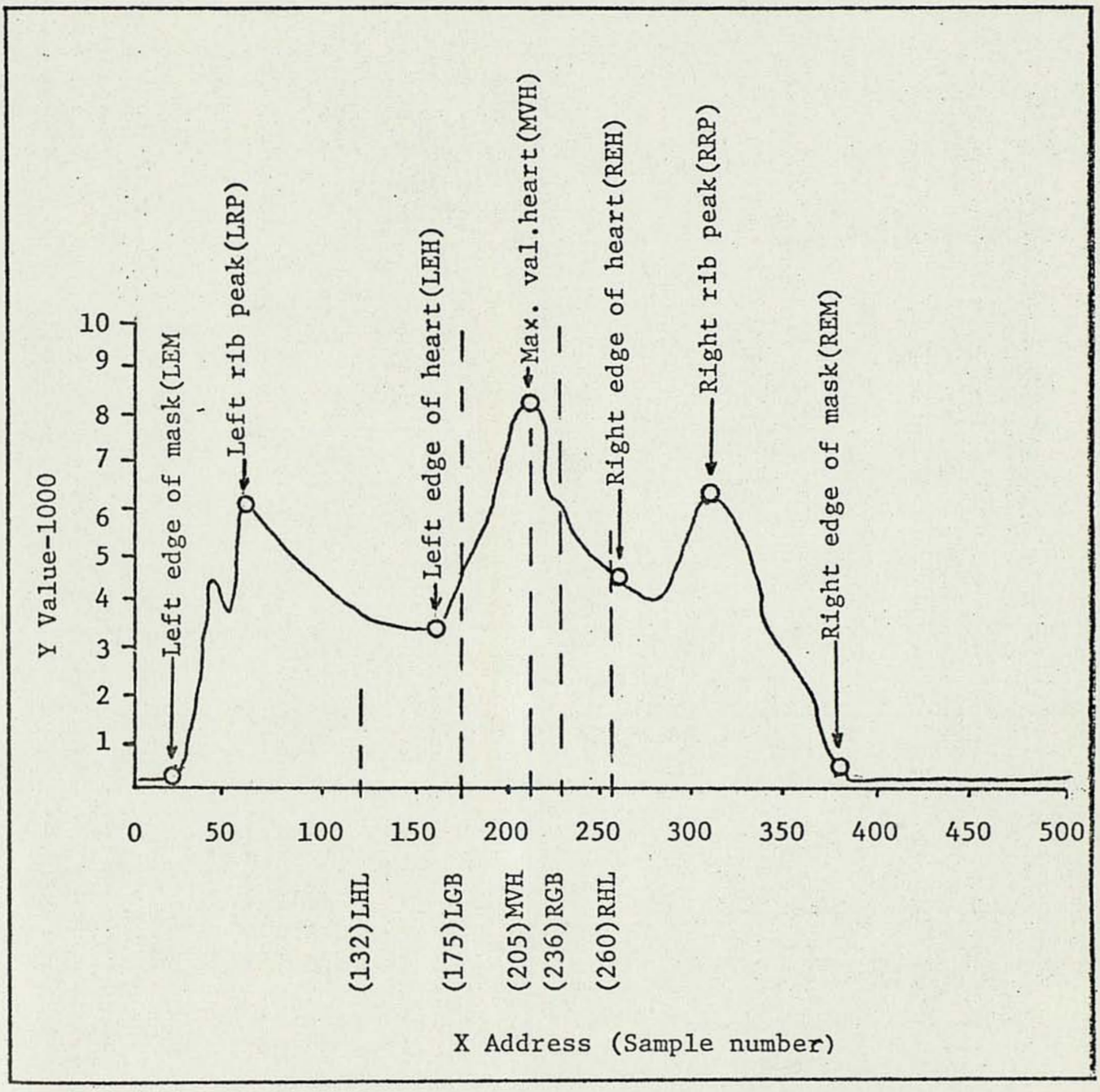

Figure 4. Typical "Sums Tape" record showing integrated Y values at each X address.

By orienting the origin point (0,0 - see Figure 4) of the Cartesian coordinate system with the zero address of the columns, we can move across the abscissa (X address) to a point (about 45 on the

X axis) where the window opens and then, all the way across the abscissa (X address) to where the window closes (about 370 on the X axis). The flying spot scanner sweep (mechanical arm) and the tape write time (response time of the mechanical arm) considerations are the reasons the images appear (the amplitude is not zero) to the left of the X address, 45, and to the right of X address 370 on the X axis.[31]

The computation of the information which has been stored to determine the cardiothoracic ratio was accomplished in two phases (IBM 1410). Phase one integrates each of the 500 columns of the matrix and stores this information as was previously discussed (Figure 4); this is considered the "Sums Tape." A condensed flow chart for this phase is shown in Figure 5.[32] The program bypasses the shoulders. "Each Y value is summed on each of the X addresses until five of these Y values at X address '131' have been equal to or greater than the number '45'." This indicates that the diaphram has been reached (under normal conditions) and the "Sums Tape" record is written.

Phase two uses the data collected in phase one and a stored computer program to determine the following points on the X address:

Left Rib Peak	(LRP)
Left Edge of Heart	(LEH)
Maximum Y Value of Heart Band	(MVH)
Right Edge of Heart	(REH)

---

[31] Ibid., p. 68.

[32] Ibid., p. 69.

Right Rib Peak	(RRP)
Maximum Transverse Diameter of Heart	(MDH)
Maximum Transverse Diameter of Thorax	(MDT)
Cardiothoracic Ratio	(MDH/MDT = CR)

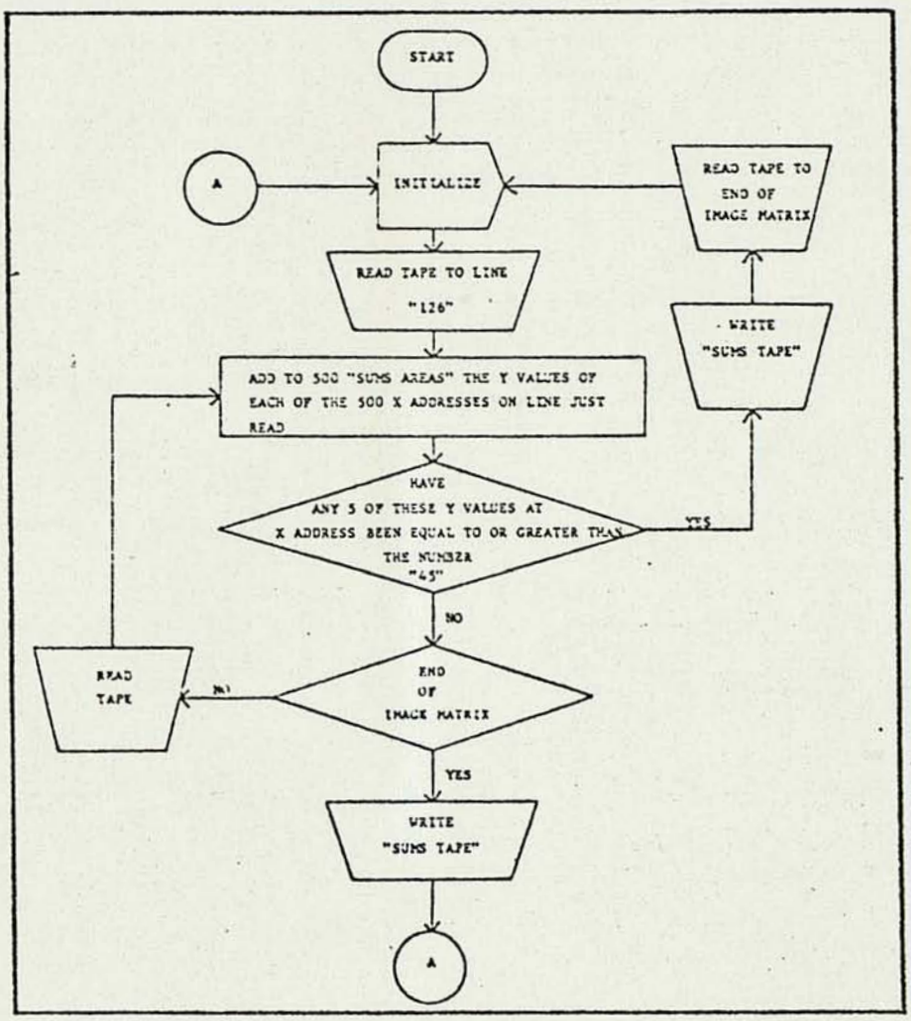

Figure 5. Phase 1, program flow chart (condensed).

A condensed flow diagram of the program is shown in Figure 6.[33]

---

[33] Ibid.

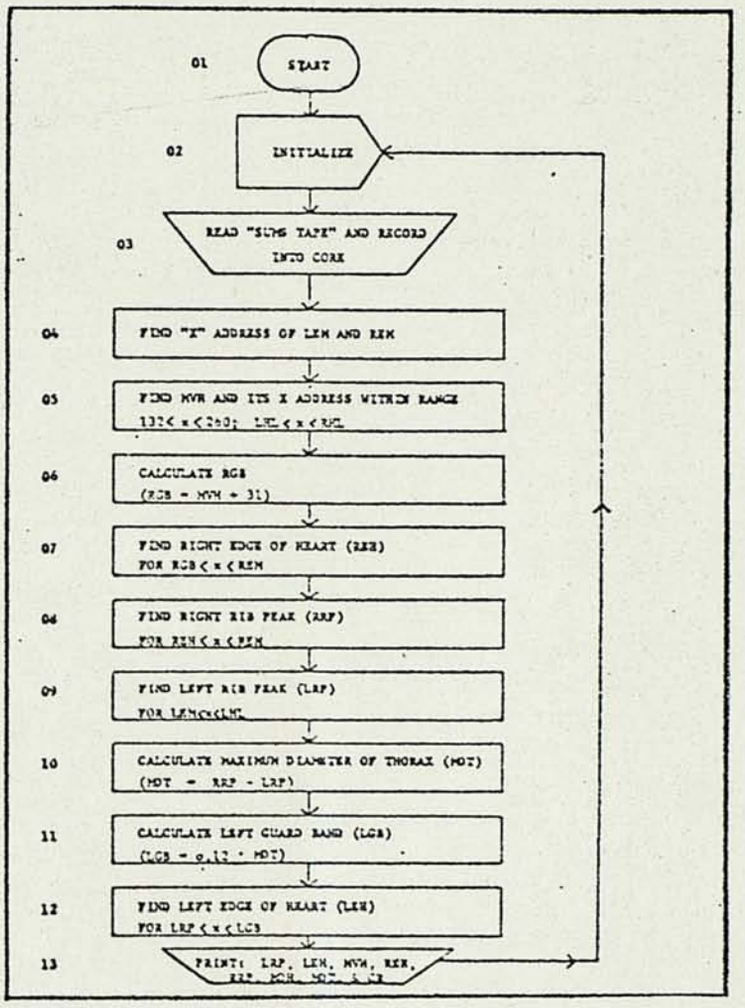

Figure 6. Phase 2, program flow chart (condensed).

In order to obtain an understanding of the procedure the program uses to locate a typical X address, the algorithm for locating the right edge of the heart (REH) is given below. The words of the author are used to preserve the meaning.

> To find the right edge of the heart, the computer searches from left to right (see Figure 4) for the X address at which the first-order difference goes positive. (If the "Sums Tape" record were a continuous curve, this would be the point where the first derivative dy/dx went positive.) This search is begun at the X address (RGB) which is the right-hand limit of a central "guard band," and it is continued to the right until a positive first-order difference is encountered. Beginning at this X

> address, a search is made from right to left over an interval whose length is defined as 10 percent of the distance between this X address and the X address of the maximum Y value of the heart peak (MVH). The X address of the maximum value off the second-order difference (analogous to the "second derivative") in this interval is taken as the locus of the right edge of the heart shadow. The guard band mentioned above is provided to insure that the search for the right and left edges of the heart are begun outside of the central region where the bony structure of the spine causes rapid changes in the Y values as a function of the X address.[34]

The results of this study are summarized in Table VI and Table VII.

Table VI is a list of cardiothoracic ratios measured by four investigators which are used to compare the accuracy of the software.[35] The four investigators were as follows: CR (a) - a Board Certified radiologist, CR (b) - a resident in radiology, CR (c) - a computer scientist, and CR (d) - a biomedical engineer.

Table VII shows a comparison of the cardiothoracic ratio calculated by the computer CR (c) and the average cardiothoracic ratios measured by the four investigators by using a percentage difference between the two.[36] Forty-five percent of the ratios were within ± 5 percent of each other. The rest, except for one, were within ± 15 percent. This one case was blamed on abnormal amounts of air in the stomach of the patient.

---

[34] Ibid., p. 70.

[35] Ibid.

[36] Ibid., p. 71.

TABLE VI

Cardiothoracic Ratios Measured by Investigators: CR(a) by a Board Certified Radiologist, CR(b) by a Resident in Radiology, CR(c) by a Computer Scientist, and CR(d) by a Bio-Medical Engineer.

CR(a)	CR(b)	CR(c)	CR(d)
0.402	0.396	0.402	0.400
0.467	0.457	0.457	0.445
0.488	0.500	0.472	0.445
0.435	0.426	0.435	0.404
0.433	0.408	0.438	0.408
0.442	0.438	0.447	0.417
0.465	0.455	0.465	0.437
0.500	0.500	0.489	0.468
0.533	0.556	0.544	0.517
0.494	0.477	0.500	0.478
0.443	0.429	0.439	0.428
0.535	0.489	0.535	0.506
0.400	0.400	0.382	0.365
0.340	0.346	0.340	0.327
0.511	0.511	0.495	0.463
0.427	0.409	0.385	0.389
0.475	0.439	0.434	0.414
0.451	0.415	0.410	0.415
0.480	0.453	0.442	0.428
0.410	0.410	0.412	0.405
0.467	0.467	0.478	0.440
0.488	0.500	0.482	0.447
0.386	0.366	0.376	0.365
0.435	0.396	0.421	0.411
0.385	0.385	0.368	0.355
0.430	0.396	0.420	0.414
0.398	0.404	0.400	0.384
0.322	0.326	0.316	0.309
0.447	0.456	0.427	0.412
0.450	0.444	0.434	0.410
0.444	0.444	0.435	0.419
0.521	0.532	0.510	0.500
0.454	0.447	0.422	0.416
0.374	0.348	0.337	0.319
0.411	0.434	0.418	0.400
0.392	0.392	0.381	0.362
0.578	0.565	0.532	0.538

TABLE VII

Cardiothoracic Ratios Calculated by the Computer CR(c), Average of CR(a-d), and Percent Difference Between CR(c) and CR(a-d)

	CR(c)	Average CR(a-d)	Percent
	0.382	0.400	- 4.5
	0.300	0.456	-34.2
	0.456	0.476	- 4.2
	0.392	0.425	- 7.7
	0.412	0.422	- 2.3
	0.392	0.436	-10.0
	0.403	0.456	-11.6
	0.416	0.489	-14.9
	0.496	0.538	- 7.8
	0.491	0.487	0.8
	0.460	0.435	5.4
	0.475	0.516	- 7.9
	0.332	0.387	-14.2
	0.343	0.338	1.4
	0.490	0.495	- 1.0
	0.400	0.403	- 0.7
	0.414	0.441	- 6.1
(a)	0.395	0.423	- 6.6
	0.466	0.451	3.2
	0.441	0.409	7.2
	0.397	0.463	-14.2
	0.472	0.479	- 1.4
	0.377	0.373	1.0
(b)	0.412	0.416	- 0.9
	0.390	0.373	4.3
	0.454	0.415	8.5
	0.429	0.397	7.4
	0.330	0.318	3.6
	0.375	0.436	-13.9
	0.441	0.434	1.5
	0.437	0.436	0.2
	0.453	0.516	-12.2
	0.416	0.435	- 4.3
	0.353	0.344	2.5
	0.444	0.416	6.3
	0.362	0.382	- 5.2
	0.491	0.553	-11.2

B. University of Missouri Study

A University of Missouri team devised a computer program that is capable of diagnosing certain rheumatic heart diseases on the basis of a patient's x-rays. This method involves extraction of an image to a two dimensional array of numbers which represent the image. These arrays of numbers are converted into meaningful measurements which can be used to diagnose according to probabilistic logic which is built into the computer program.[37]

The two dimensional array of numbers represents the grayness of the image or, better yet, the graduation of light and dark at each particular location on the film. The camera used in this particular study was capable of distinguishing 64 shades of gray and can resolve an image into an array as large as 1024 by 1024 at a rate of 1000 points per second. The application in this study to rheumatic heart diseases only used an array of 256 by 256.[38]

The most difficult part of this process is differentiating the grayness values in order to identify and extract the features due to the overlapping organs such as the heart, lungs, and rib cage. It was pointed out that several empirical techniques evolved to measure the side and shape of the heart which are important in the diagnosis of

---

[37]"The Computer as a Doctor," Newsweek, December 27, 1971, p. 52.

[38]Allen L. Hammond, "Image Analysis: Application to Automated Medical Diagnosis," Science, December 3, 1971, p. 1011.

rheumatic heart disease. (The heart is characteristically enlarged or disfigured in persons having rheumatic heart disease.)[39]

Approximate positions of the major organs (heart, lungs) are found by scanning the regions whose intensities are roughly continuous and whose size and average gray values are close to those expected.

These approximate positions are tested by the computer by assuming, for example, the physiological reality of the position of the right lung to the right side of the heart.

Once organs are identified, points bordering the organs are examined so that points identifying its outline can be appropriately assigned. The overlapping of the different regions within the image allows the boundaries of the major organs to be determined. The program can then outline the heart and extract information which allows a diagnosis to be made. "Curves (fourth order polynomials) are fitted to the left and right boundaries of the heart, horizontal and diagonal dimensions are measured, and the ratio of heart to lung area is determined. A total of thirteen measurements are made."[40]

The technique described above of measuring a ratio of the heart to lung areas was tested against a panel of ten radiologists making the same measurements for 135 cases. The correct diagnoses of the cases were known prior to the test but the results were not shown to the radiologists. The radiologists were allowed to use both frontal and lateral x-rays while making individual diagnoses, while the program

---

[39]"The Computer as a Doctor," p. 53.

[40]Hammond, "Image Analysis," p. 1011.

used only the frontal x-rays. The computer's accuracy was 73 percent, 11 percent higher than the overall accuracy of the ten radiologists and also higher than any of the individual diagnoses.

The authors point out that Lodwick's view was that "the performance of the image analysis system and the accuracy of the fully automated diagnosis thus established the feasibility of using image analysis for this medical application."

At the present time clinical evaluation of this system is underway at the University of Missouri, with hopes of continued positive results.[41]

2. Computer Diagnosis in Chromosome Analysis

Robert S. Ledley presented a study to show the possible aid of computers in the diagnosis of disease characterized by chromosomal disorders. Photomicrographs of cells are made in one of the stages during cell division (metaphase stage), at which time the chromosomes are well distinguished.

An electronic scanning instrument called the Film Input to Digital Automatic Computer (FIDAC) is used to make a photomicrograph transparency of a picture pattern, a chromosome spread in this instance by reading a number which represents a numeric gray level density of the transparency into the computer's memory.[42] This numeric gray-level density, a number which represents the possible

---

[41] Ibid., p. 1012.

[42] Robert S. Ledley, "Computer Aids to Medical Diagnosis," Journal of the American Medical Association, CXCVI (June 13, 1966), 933-43.

shades of gray between transparency (clear) and darkness (black), is measured at each point on the picture at a rate of about 350,000 points per .3 seconds.[43]

Once the FIDAC has read the picture pattern into the computer (IBM 7094), the program recognizes individual chromosomes by using a pattern recognition analysis comprising four phases. The first phase of the program involves the proper placement of the picture in the computer's memory. It may be desirable from time to time to change the size of the picture for clarity and more detailed analysis. The second phase of the picture analysis program involves a "parts analysis" of the object in the picture being considered in which the computer scans the picture until an object is found by a significant change in the shades of gray (dark spots). At this point, the "bug," a program that causes a scan to move systematically across the picture looking for a dark spot, is programmed to follow the boundary of the dark spot. This program characterizes a list of different segments by their average direction and general curvature. The third phase of the computer program involves identifying the segments found (listed in a Boundary List) in the second phase with "parts," i.e., predefined segments. As each set of segments from the boundary list satisfies one of the "parts definitions, the part is considered identified, and the identification is recorded in a 'parts list.'" The parts list is retained in the order which they occurred to preserve the outline of the object. The fourth phase of the picture analysis program involves

---

[43] Ibid.

analysis of the parts list for certain dimensions.[44] The area and length of a chromosome is valuable information for diagnosis, and when determined by the computer, takes less than 1/2 second for each chromosome.[45]

The author pointed out that the time reduced by using this computer technique to analyze the chromosomes saves hours of detailed manual analysis. Ledley also states that some of the measurements that are made by the computer, such as areas and integrated densities, are difficult to make accurately by manual means.

3. Computer Diagnosis in Differential Preoperative Diagnosis for Pelvic Surgery

Neurath, Enslein and Mitchell designed a program to assist surgeons in arriving at a preoperative diagnosis for pelvic surgery. The program collects information by means of a checklist. The information for the checklist is coded and key punched on two computer cards, which are then compiled and typed onto a medical record to be used by the surgeon. The program also computes the most probably diagnosis and disease probability on the basis of the information submitted, and prepares a typed list of the diagnosis, together with the prevalence, in this group of patients. If the program finds than an item is missing on the checklist, a request is made to obtain this information in a subsequent run.

The data used in this checklist was extracted from information

---

[44]Robert S. Ledley, Use of Computers in Biology and Medicine (New York: McGraw-Hill Book Company, Inc., 1965), p. 482.

[45]Ledley, "Computer Aids," p. 943.

collected by three teaching hospitals and includes information which has been agreed upon by a number of consultants as being relevant and complete in the field of pelvic surgery. An example of the first page of the current checklist is shown in Table VIII.[46] This information was classified in different categories such as age, marital status, etc., numbered numerically in order of "severity." The numerical value assigned to each item is a weighted value which reflects the severity of each item within its category. The larger the number the greater its severity. For example, the first category, age, the age group number one (11-20) was less critical or severe when undergoing pelvic surgery than age group number 9 (66+). The number one is less in numeric value than the number 9 and therefore is less severe.

An average of about 35 percent of the information requested on the medical record (checklist) was not available. The missing information was assigned a value in a normal range in order to complete the study. It was felt that this information would have been present if the symptoms (any information about the patient) were abnormal.

Information from over 500 patients was originally coded with nine diagnoses. Of the approximately 500 cases, seventy-five were eliminated from the study due to complications.[47] Table IX shows how the 425 cases were divided among the nine diagnoses.[48]

---

[46]Peter W. Neurath, Kurt Enslein, and George W. Mitchell, Jr. "Design of a Computer System to Assist in Differential Preoperative Diagnosis for Pelvic Surgery," The New England Journal of Medicine, CCLXXX (April 3, 1969), 745.

[47]Ibid., p. 747.

[48]Ibid.

TABLE VIII

First Page of Checklist

PELVIC SURGERY

Patient's Name: _____

Date Admitted: _____
Month Day Year

NECH PATIENT NO. 1-6 [ ] SURGEON'S NAME 7-17 [ ]

Form Completed By: _____

DIRECTIONS: Please circle number under *EACH* category.

A. HISTORY

AGE: col. 18
1. 11-20
2. 21-30
3. 31-35
4. 36-40
5. 41-45
6. 46-50
7. 51-55
8. 56-65
9. 66+

MARITAL STATUS: col. 19
1. Not married
2. Married
3. Separated
4. Divorced
5. Widowed

COITUS: col. 20
1. No
2. Yes, with reliable contraception
3. Yes

MENOPAUSE: col. 21
1. >1 yr
2. <1 yr
3. Premenopause

HORMONES: col. 22
1. None
2. Estrogen
3. Progesterone
4. Estrogen & progesterone
8. Thyroid & testosterone

AMENORRHEA: col. 23
(other than postmenopause)
1. None
2. Primary
3. Acquired

INTERVAL SINCE LMP: col. 24
0. Not menstruating at all
1. Normal
2. Occurred much sooner than usual
3. Delayed >1 wk, but <1 mo
4. Delayed >1 mo

AVERAGE DURATION OF MENSTRUAL PERIOD: col. 25
(If postmenopause, give premenstrual values)
1. 1-2 days
2. 3-7 days
3. 8-10 days
4. 11+ days

AVERAGE MENSTRUAL FLOW col. 26
(If postmenopause, give premenopause values)
1. Normal (includes scant or heavy)
2. Profuse, clots
(If 2):
DURATION OF SUCH FLOW col. 27
1. <30 days
4. 31-90 days
5. 91-364 days
7. 1-4 yr
9. 4+ yr

SPOTTING: col. 28
1. None
2. Mid-cycle
3. Postmenstrual
4. Continual
(If 2,3,4):

DURATION OF SUCH SPOTTING: col. 29
1. 30 days or less
2. 31 days or more

LOW BACKACHE: col. 30
1. None
2. With menses
3. Always

DYSMENORRHEA: col. 31
1. None
2. Primary
3. Acquired < or = 1 yr
4. Acquired >1 yr

LOWER ABDOMINAL PAIN col. 32
1. None
2. LLQ
3. RLQ
4. Bilateral
(If 2,3 or 4):
DURATION OF LOWER ABDOMINAL PAIN: col. 33
1. 30 days or less
2. 31 days or more

VAGINAL DISCHARGE col. 34
1. None
2. Yes
3. With pruritus
4. Bloody
(If 2,3 or 4):
DURATION OF VAGINAL DISCHARGE col. 35
1. 30 days or less
2. 31 days or more

The mathematical model to analyze the information collected is described by the authors as follows:

> The diagnostic method employed by our system is a statistical-pattern recognition method. It uses a discriminant function, which is obtained from the data described by a so-called stepwise linear discriminant analysis. For this application the method has notable advantages over others, particularly from the point of view of ease of implementation.[49]

---

[49] Ibid.

The method described above can be explained easily in a two dimensional situation where there are two possible diagnoses, A and B, for a patient and two possible symptoms, X and Y. If the two symptoms are plotted on an X-Y cartesian coordinate system for a number of patients, a difference between the symptoms A and B can be seen as shown in Figure 7.[50]

TABLE IX

Accuracy of Diagnosis in 425 Patients (+ Represent Range from 19% to 100% - Only Average Error Rates Were Known)

Pathologist's Final Diagnosis	No. of Cases	% Incorrect Preoperative Diagnosis	
		Surgeon	Computer
Fibroid	144	+	35
Stein - Leventhal syndrome	9	+	44
Endometriosis	18	+	22
Ovarian tumor	49	+	35
Prolapse	119	2	29
Endometrial polyp	35	+	43
Ectopic pregnancy	16	+	25
Pelvic inflammatory disease	16	+	44
Cancer	9	+	33
All diagnoses	425	34	33

The ordinate axis represents the severity of symptom Y, which increases as you move up, and the abscissa axis represents the severity of symptom X, which increases as you move to the right. Several cases have been plotted with patients diagnosed with disease A

[50] Ibid.

represented by "X" and those with disease B represented by "0." The diseases of the patients in this instance are represented by a cluster of points of which the mean value has been designated by $(\overline{X}_A \overline{Y}_A)$ and $(\overline{X}_B \overline{Y}_B)$. An attempt is now made to separate the two types of points representing diseases of patients. The best straight line (for a two dimensional plane) can be found by statistical rules and, as is shown in Figure 7, separates the points rather well. The probability that a new patient's diagnosis will be A or B with symptoms X and Y can be determined by the relationship of the point to the line.

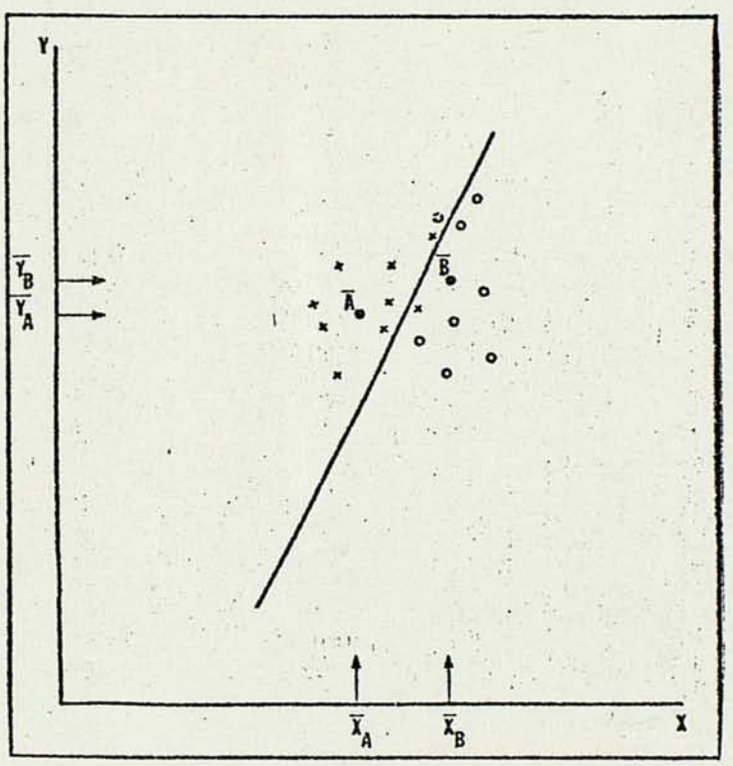

Figure 7. Graph of the two-system combination to differentiate two diagnoses in a large group of patients.

The technique described above can be used in three dimensions with a plane separating two clusters for any two diagnoses, and for N

dimensions with a hyperplane of dimension (N-1) separating the clusters. Four or more dimensions may be difficult to visualize, but a set of equations might help to explain the feasibility of the technique. The equations which are to follow are part of the algorithm available from the University of California (Los Angeles) library of statistical routines. Up to 100 different variables (symptoms) can be processed by this software.

With the information derived from the 425 cases with which we are dealing, fifty-three symptoms were considered with all the patients having one of the nine diseases shown in Table IX. The program selected the most discriminating symptoms, thus eliminating several symptoms and leaving twenty-six. Thus the equation relating the nine diseases and the twenty-six symptoms looks as follows.

$$D_1 = B_{10} + B_{11}X_{11} + \cdots + B_{in}X_{in}$$
$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots$$
$$D_i = B_{i0} + B_{i1} + \cdots + B_{ij}X_{ij} \cdots + B_{in}X_{in}$$
$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots$$
$$D_9 = B_{90} + \cdots + B_{926}X_{926}$$

The B_{ij}s represent the coefficients (weights) and the X_{ij}s represent the average severity of the "j^{th}" symptom for patients with the diagnosis of "i." In the equations above, the twenty-six most discriminant systems have been identified in Table X, ranked on the basis of discriminant power.[51]

The final step of the method is to determine the diagnosis of any patient by means of the discriminant functions. The symptoms

---

[51] Ibid., p. 748.

are substituted into the nine equations representing the nine diseases possible, with the largest value of D_i as the most probable diagnosis. The error rate of the computer and of the surgeons is shown in Table IX. The two error rates were almost equal.

TABLE X

Rank Order of 26 Most Discriminating Factors

Rank	Name	"F" Posterior
1	Miscellaneous symptoms	18.52
2	Age	9.09
3	Prolapse	8.64
4	Condition	8.53
5	Adnexa	6.56
6	Pelvic masses	6.01
7	Previous examination	5.63
8	Urinary symptoms	5.35
9	Dysmenorrhea	5.08
10	Abdominal x-ray studies - before operation	5.05
11	Pregnancy test	4.66
12	Previous pelvic operations	3.94
13	Duration of irregular flow	3.85
14	History of acute pelvic inflammatory disease	3.79
15	Uterus	3.68
16	Pain in lower abdomen	3.30
17	Neck	3.30
18	Sterility	3.13
19	Bearing down	3.13
20	Hormones	3.06
21	Hemoglobin before operation	2.91
22	Biopsy of smear	2.55
23	Last menstrual period	2.24
24	Tenderness	1.94
25	Flow	1.94
26	Microscopical examination of urinary sediment - before operation	1.84

One point the authors make clear concerning the conclusiveness of the tests is that the discriminant function used was constructed and tested on the same patients' data from which it was derived. They therefore consider it not as conclusive a test as it could be if a new set of data from other patients were considered. However, they do remark that the results appear good enough to justify establishing a working system based on the method and the data base described.[52]

The technique seems relatively sound but its application cannot be accepted without first being tested on many patients, allowing enough time to follow up its efficiency.

There are several long and short range objectives which the authors have presented. They are as follows:

A. Long range objectives -

1. Establishment of a better data base.
2. Improvement of the discriminant function and testing of its error rates against new but compatible data.
3. Determination of the minimum list of symptoms needed to make a differential diagnosis of desirable accuracy.

B. Short range objectives -

1. Information about the patient of immediate usefulness to the operating surgeon.
2. Copies of a neat, computer-typed medical history for the surgeon and his associates within one hour of submission of the checklist.
3. A computed differential diagnosis, printed and delivered with the history.[53]

---

[52] Ibid.

[53] Ibid., pp. 748-49.

4. Computer Diagnosis in Evaluation of Electrolyte and Acid-Base Disorders

Howard L. Bleich translated the current concepts of acid-base disorders into a time-sharing computer program which can be used by physicians in diagnosing patients' disorders by terminal connection to the Bell Telephone System. As of January 1971, private telephone wires were extended to Washington, D.C., New York City, Detroit and Chicago from the computer center in Cambridge, Massachusetts.[54] The technique used by the digital equipment corporation computer (Model PDP-1D) permits its services to be used by sixty-four remotely located teletype terminals simultaneously, with almost instantaneous response.[55]

An example of the dialogue which takes place between the user at the terminal and the computer is seen in Figure 8.[56] The program, once called, prints the title, date and time, and then requests information concerning serum electrolytes starting with "NA = ___," pausing to allow the user to enter the information in the blank before going to the next line, "K = ___." If the information requested is not available the user pushes an "enter" button. It is possible for the program to calculate values that were not supplied by the physician.

---

[54]Howard L. Bleich, "The Computer as a Consultant," Seminars in Medicine of the Beth Israel Hospital (Boston), CCLXXXIV (January 21, 1971), 143.

[55]Howard L. Bleich, "Computer Evaluation of Acid-Base Disorders," Journal of Clinical Investigation, XLVIII (September, 1969), 1691.

[56]Ibid., p. 1692.

```
ACID-BASE EVALUATION   MAY 15, 1969     5:46 PM


SERUM ELECTROLYTES (MEQ/L):

   NA      =  140
   K       =  4.0
   K       =  3.9
   CL      =  110
   CO2T    =   20

BLOOD PH   =  7.19

BLOOD PCO2 =  49.9  (CALCULATED FROM CO2T AND PH)

PATIENT'S WEIGHT (IN POUNDS)      =  150

IS THERE EVIDENCE OF PULMONARY CONGESTION OR CONGESTIVE HEART FAILURE?
NO

                    EVALUATION NOTE

     THE PH OF 7.19 UNITS IS LOWER THAN CAN BE ACCOUNTED FOR
BY THIS DEGREE OF HYPERCAPNIA, AND INDICATES THAT A METABOLIC
ACIDOSIS IS SUPERIMPOSED ON THE RESPIRATORY ACIDOSIS.

     FURTHERMORE, THE FINDING OF A NORMAL PLASMA CONCEN-
TRATION OF UNMEASURED ANIONS (14.0 MEQ/L) INDICATES THAT THE MOST
LIKELY CAUSES OF THE METABOLIC ACIDOSIS ARE:
   1) PYELONEPHRITIS WITHOUT GLOMERULAR FAILURE
   2) RENAL TUBULAR ACIDOSIS
   3) CARBONIC ANHYDRASE INHIBITION (DIAMOX)
   4) AMMONIUM CHLORIDE INGESTION
   5) URETEROSIGMOIDOSTOMY
   6) MASSIVE DIARRHEA, PANCREATIC DRAINAGE, OR
      FISTULAE OF THE SMALL INTESTINE.

     IN AN EFFORT TO CORRECT THE METABOLIC COMPONENT OF THE ACIDOSIS
IT IS SUGGESTED THAT SUFFICIENT ALKALI BE GIVEN TO RAISE PLASMA
BICARBONATE CONCENTRATION TO A VALUE THAT WOULD BE MORE IN KEEPING
WITH UNCOMPLICATED HYPERCAPNIA.

THE CALCULATED QUANTITY OF BICARBONATE REQUIRED TO ACHIEVE THIS GOAL
IS APPROXIMATELY 250 MEQ.  IT IS SUGGESTED, HOWEVER, THAT 132 MEQ
OF SODIUM BICARBONATE BE GIVEN DURING THE NEXT FEW HOURS, AND THAT
SERUM ELECTROLYTES AND BLOOD PH BE MEASURED AGAIN AT THAT TIME.

     FINALLY, EVERY EFFORT SHOULD BE MADE TO IMPROVE PULMONARY
VENTILATION AND TO MAINTAIN BLOOD PCO2 AT A VALUE OF NO GREATER
THAN 45 MM HG.

     THANK YOU FOR REFERRING THIS INTERESTING PATIENT TO US.

REFERENCES:

1) BRACKETT, N.C., JR., COHEN, J.J., AND SCHWARTZ, W.B.
   CARBON DIOXIDE TITRATION CURVE OF NORMAL MAN.
   NEW ENG. J. MED., 272, 6, 1965.

2) BESSON, P.B. AND MCDERMOTT, W.
   CECIL-LOEB TEXTBOOK OF MEDICINE
   10TH EDITION (1967), P. 762.
```

Figure 8. A sample acid-base time-sharing program output.

If the user is aware of a mistake he may erase it by means of a "rub out" button. The program checks the compatibility of the information entered so that when mistakes are found, correct entries are again requested as has been done for the value of "K" in Figure 8. (If work on a problem begins but is not completed when an allotted time interval has elapsed, the unfinished program is switched out.)[57]

After the last entry the computer prints diagnostic notes (as can be seen in Figure 8), the length of which may be as short as four words or as long as two pages, varying with the complexity of the abnormality presented. The program suggests possible disorders and possible therapy, indicates determinations that should be remeasured, and provides a list of possible medical references.[58]

Figure 9 shows a specific case in which insufficient information was supplied.[59] When information is not available the program is designed to draw the most useful conclusions possible. In this instance, the computer determined values for the blood Ph and blood PCO2. As the data from the patient approaches the border zones where the confidence with which a particular diagnosis can be made decreases, modifiers such as "probably," and "relatively" are inserted into the text.

---

[57]Bleich, "The Computer as a Consultant," p. 143.

[58]Bleich, "Computer Evaluation of Acid-Base Disorders," p. 1693.

[59]Ibid., p. 1694.

```
ACID-BASE EVALUATION   MAY 15, 1969   5:51 PM

SERUM ELECTROLYTES (MEQ/L):

   NA    =

   CO2T  =  36

BLOOD PH  =

BLOOD PCO2 =

IN THE ABSENCE OF A MEASURE OF BLOOD PH OR PCO2 WE CANNOT DETERMINE
WHETHER THE ELEVATED PLASMA BICARBONATE CONCENTRATION IS DUE TO
RESPIRATORY ACIDOSIS, TO METABOLIC ALKALOSIS, OR TO A MIXED ACID-BASE
DISORDER.
DO YOU WISH TO:
   1) ASSUME THAT THE DISORDER IS ENTIRELY RESPIRATORY
   2) ASSUME THAT THE DISORDER IS ENTIRELY METABOLIC
   3) SUPPLY THE COMPUTER WITH BLOOD PH OR BLOOD PCO2
PLEASE SELECT A NUMBER>...   2

PATIENT'S WEIGHT (IN POUNDS)   =  140

DO YOU HAVE A COMPLETE SET OF SERUM ELECTROLYTES?>...   NO

IT WOULD PROBABLY BE USEFUL TO OBTAIN THEM.


     IF WE ASSUME THAT THE ACID-BASE DISTURBANCE IS ENTIRELY METABOLIC,
THE CO2T OF 36 MEQ/L IMPLIES THAT PH MUST BE 7.43 UNITS AND
PCO2 46 MM HG. SINCE THESE ARE THE VALUES THAT WOULD BE COMPAT-
IBLE WITH UNCOMPLICATED METABOLIC ALKALOSIS.

     WHILE AWAITING A COMPLETE SET OF ELECTROLYTES IT SHOULD BE
RECALLED THAT METABOLIC ALKALOSIS IS MOST COMMONLY CAUSED BY:
   1) VOMITING OR GASTRIC ASPIRATION
   2) DIURETIC THERAPY (THIAZIDES, MERCURIALS, ETHACRYNIC ACID, OR
      FUROSEMIDE)
   3) HYPERADRENALISM (CUSHINGS SYNDROME, PRIMARY OR SECONDARY
      ALDOSTERONISM, ADRENAL STEROID THERAPY)
   4) EXCESSIVE ALKALI INTAKE (ALKALOSIS USUALLY TRANSIENT)
   5) POST HYPERCAPNIA (PARTICULARLY IF CHLORIDE INTAKE IS SMALL)
   6) HYPERCALCEMIA (POSSIBLE--DATA NOT CLEAR)
   7) BARTTER'S SYNDROME (HYPOKALEMIA, HYPOCHLOREMIA, AND METABOLIC
      ALKALOSIS, WITH, IN SOME INSTANCES, HYPONATREMIA, SHORT STATURE,
      AND MENTAL RETARDATION.)

     IF HYPERADRENALISM AND HYPERCALCEMIA ARE ABSENT, THEN IT SHOULD
BE POSSIBLE TO CORRECT THE METABOLIC ALKALOSIS BY ADMINISTERING
SUFFICIENT CHLORIDE TO REPLACE PREVIOUS LOSSES AND TO ALLOW A
URINARY CHLORIDE EXCRETION OF AT LEAST 10 TO 20 MEQ PER DAY.

THE TOTAL CALCULATED CHLORIDE DEFICIT IS APPROXIMATELY 130 MEQ.
IT IS SUGGESTED, HOWEVER, THAT IN ADDITION TO REPLACING KNOWN CHLORIDE
LOSSES, 100 MEQ OF SODIUM, POTASSIUM, OR ARGININE CHLORIDE BE
GIVEN INTRAVENOUSLY DURING THE NEXT 24 HOURS, AND THAT SERUM ELECTRO-
LYTES AND BLOOD PH BE MEASURED AT LEAST DAILY UNTIL SIGNIFICANT
IMPROVEMENT IN THE ACID-BASE ABNORMALITY OCCURS.

     THANK YOU FOR REFERRING THIS INTERESTING PATIENT TO US.

REFERENCES:

1) GOLDING, E.M., CANON, R.J., HEINEMANN, H.O., AND FISHMAN, A.P.
   RESPIRATORY ADJUSTMENT TO CHRONIC METABOLIC ALKALOSIS IN MAN
   J. CLIN. INVEST., 47, 155, 1968.

2) PEARCE, P.D. AND MODERNATT, W.
   CHILD-LOVED TENDENCIES OF MEDICINE
   12TH EDITION (1967), P. 763.

3) HARRISON, J.F., FISHMAN, R.H., LAWRENCE, D.R., AND SUNSHINE, M.H.
   THE OPTIMAL ROLE OF CHLORIDE IN THE CORRECTION OF HYPOKALEMIC
   ALKALOSIS IN MAN.
   [illegible] 1965.
```

Figure 9. A sample acid-base time-sharing program output.

Figure 10 shows a portion of the logic of the program which begins by collecting data.[60] As can be seen in the figure, the program checks the input data for compatibility at several steps. For example, the program requests values for Na^+, K^+, Cl^-, and CO_2T. Unless the value of $[Na^+ + K^+ - Cl^- - CO_2T]$ is greater than 6, the program realizes that one or more of the values are incorrect and prints comment explaining incapatibility. The program is also designed to calculate missing values. An example of this process can be seen in the diamond shaped step at the first right hand path (excluding the feedback loop) of the flow chart (Figure 10). The program recognizes that only two of the three values expected are given at this step and therefore the missing value must be calculated. This missing value is calculated assuming normal conditions for the patient.[61]

The author explains the language as follows:

> The program was written in MUMPS, a special, text-manipulating language developed at the Massachusetts General Hospital. Earlier versions of the program were translated into BASIC, COURSEWRITER and FORTRAN, and in these forms it is available at the Dartmouth Medical School, in hospitals served by the Ohio State Regional Medical Program and at the University of California, Davis. In most cases, however, we have discouraged efforts to translate the program; translation into the more familiar languages lacks many of the text-manipulating features required by the program, involves considerable effort, and to run the program in such languages has always proved more costly than to run it in MUMPS. More important than cost, however, is the fact that improvements continually made in the MUMPS version do not immediately become available in the translated version.[62]

---

[60] Ibid., p. 1690.

[61] Ibid., p. 1693.

[62] Bleich, "The Computer as a Consultant," pp. 143-44.

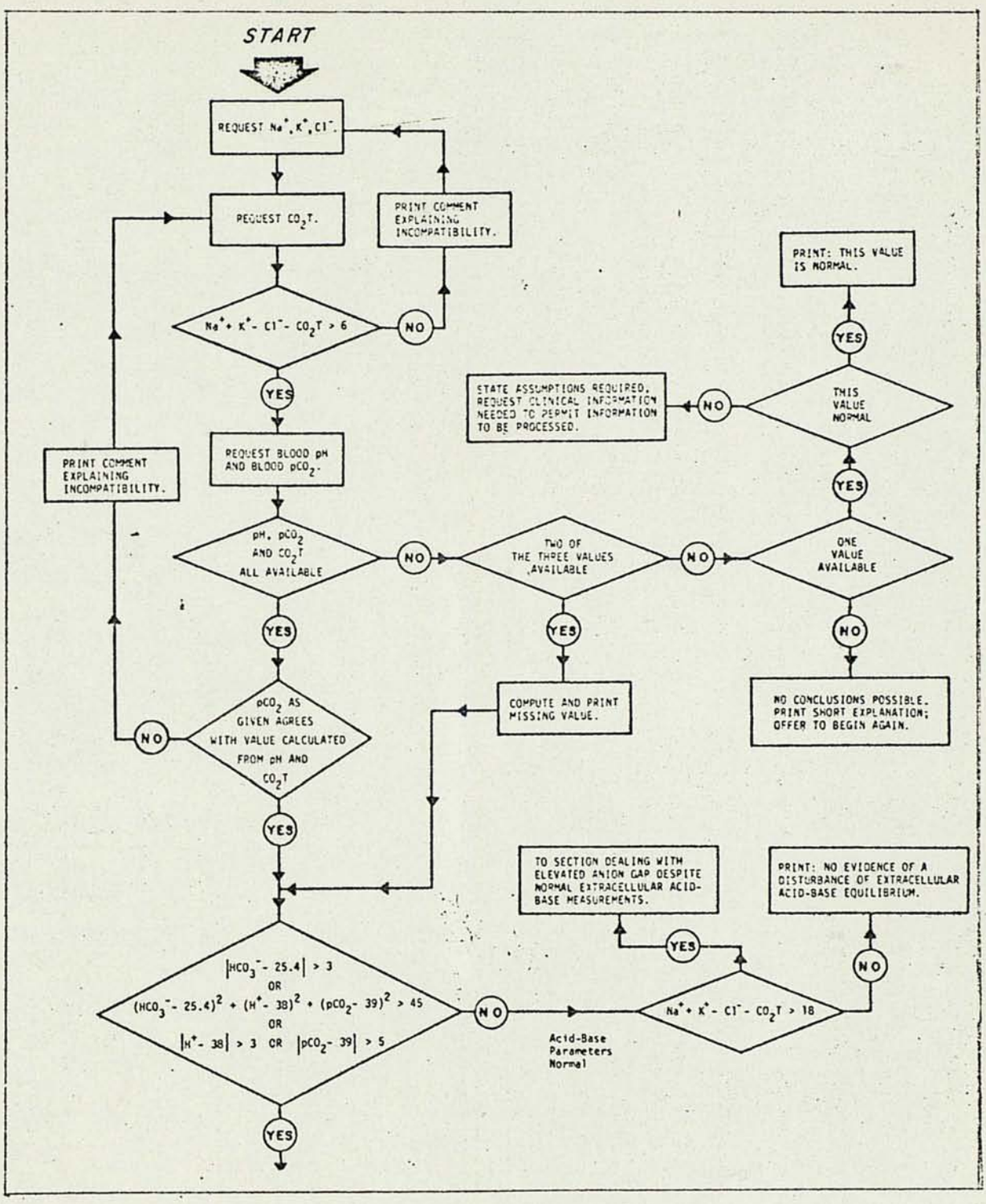

Figure 10. A sample of the acid-base time-sharing program flow chart.

No follow-up work on this subject has been published, but the author has communicated his willingness to allow others to use his program upon request.

5. Computer Diagnosis for Thyroid Disease

This application of computer diagnostic software to thyroid disease is primarily based on Bayes' Theorem, which was explained in Part I. Much of the work in this area of computer diagnosis has been done by Overall, Williams and Fitzgerald at the University of Florida. Additional work, however, has been done by Reicherz, Winkler, and Kloss at the University of Texas Medical Branch in Galveston, Texas and the University of Bonn in West Germany, by Robin, Refetoff and Selenkow at the Cambridge Hospital, Cambridge, Massachusetts and by Oddie at the University of Arkansas Medical Center. Examples of the work are presented in this section.

A. University of Florida Studies

One method developed by Williams, Fitzgerald and Overall to predict the probability of a patient having a hyperthyroid, euthyroid, or hypothyroid was to collect data concerning the patient and compare this data with a set of data established by many previous cases.[63]

Data was collected from the 1,379 cases in the format which can be seen in Table XI.[64] This information was placed in a data matrix (Table XII), which is a list of the incidence of the various signs and symptoms presented in a given category.[65] For example, by

---

[63] Clyde M. Williams, Lawrence T. Fitzgerald, and John E. Overall, Recent Advances in Diagnosis of Cancer (Chicago: Year Book Medical Publishers, 1966), p. 247.

[64] Ibid., pp. 252-54.

[65] Ibid., pp. 255-56.

referring to Table XII, it can be seen that Item 11, nervousness, was found as a symptom in 7/91 or 7.7 percent of the hypothyroid cases.

TABLE XI

Sign and Symptom Data Sheet Sample

Enter 0 if absent. Enter 1 if present. Do not enter anything if the datum is missing.

Col.	Data		Col.	Data
1-5		Patient ID Patient's Name Number		
6-7		Age	8	Sex (male 0, female 1)
9		Race (white 0, colored 1)	10	
11		Recent onset of nervousness	19	Recent onset of lethargy
12		Recent increase in heat sensitivity	20	Recent onset of cold sensitivity
13		Recent increase in sweating	21	Recent decrease in sweating
14		Recent increase in appetite	22	Recent decrease in appetite
15		Recent weight loss	23	Recent weight gain
16		Hyperkinetic movements	24	Lethargic movements
17		Warm moist skin	25	Dry coarse skin
18		Fine finger tremor	26	Facial edema
27		Eye signs (lid retraction, lag, exophthalmos)	28	Recent pain or tenderness in the thyroid gland

The theory is then based on the fact that "the probabilities of the joint occurrence of two or more independent events is the product of their separate probabilities." The actual mechanics are compiled by computer.

TABLE XII

Sign and Symptom Data Matrix (Sample)
(1,379 Cases Considered Thus Far)

	Hypothyroid		Euthyroid		Hyperthyroid	
Base Rates	.105	(145/1379)	.719	(992/1379)	.175	(242/1379)
(11) Nervousness	.077	(7/91)	.394	(297/754)	.886	(163/184)
(12) Heat sensitivity	.001	(0/82)	.185	(135/729)	.701	(115/164)
(13) Increased sweating	.102	(5/49)	.222	(155/697)	.584	(90/154)
(14) Increased appetite	.019	(1/54)	.100	(71/713)	.536	(97/181)
(15) Weight loss	.104	(7/67)	.378	(292/773)	.792	(160/202)
(16) Hyperkinetic movements	.001	(0/85)	.086	(60/701)	.670	(118/176)
(17) Warm moist skin	.001	(0/93)	.107	(75/702)	.648	(116/179)
(18) Finger tremor	.011	(1/93)	.187	(134/717)	.780	(149/191)
(19) Lethargy	.857	(78/91)	.135	(101/747)	.016	(3/185)
(20) Cold sensitivity	.786	(66/84)	.141	(103/728)	.042	(7/165)
(21) Decreased sweating	.320	(16/50)	.023	(16/698)	.001	(0/155)
(22) Decreased appetite	.434	(23/53)	.202	(144/713)	.144	(26/180)
(23) Weight gain	.477	(31/65)	.173	(133/767)	.040	(8/200)
(24) Lethargic movements	.844	(76/90)	.059	(41/694)	.023	(4/175)
(25) Dry coarse skin	.896	(86/96)	.112	(78/699)	.017	(3/178)
(26) Facial edema	.758	(72/95)	.048	(34/706)	.036	(7/194)
(27) Eye signs	.130	(7/54)	.069	(50/723)	.358	(59/165)

* Numbers in parentheses in left column refer to signs and symptoms in Table 2. A. Thus, recent onset of nervousness (11) is found in 7 out of 91 patients with hypothyroidism (7.7 per cent), etc.

An example of this theory is now presented. A patient is suspected of thyroid disease. His symptoms are warm, moist skin (ms) and increased sweating (is). The analysis would be as follows (Items 17 and 13 in Table XII):

Hypothyroid $P(S_{ms}|D_{hypo})P(S_{is}|D_{hypo}) =$

$(.001)(.102) = .000102 = .0001$

Euthyroid $P(S_{ms}|D_{euth})P(S_{is}|D_{euth}) =$

$(.107)(.222) = .023754 = .0597$

Hyperthyroid $P(S_{ms}|D_{hype})P(S_{is}|D_{hype}) =$

$(.648)(.584) = \underline{.377848} = \underline{.9402}$

Total .401704 1.0000

The right hand column shows the results normalized for convenience. Thus the probability of a hyperthyroid patient having the symptoms mentioned above is 9.402 in 10, for a euthyroid patient to have the symptoms is 0.597 in 10, and for a hypothyroid patient to have the symptoms the probability is 0.001 in 10.

If we are interested in knowing whether a patient with moist skin (ms) is likely to euthyroid, hyperthyroid or hypothyroid, it will be necessary to assign relative frequencies to each disease. Say, for example, that the relative frequency of hypothyroidism is 0.15, that for euthyroidism it is 0.65, and that for hyperthyroidism it is 0.20 in the patient population being considered.

Hypothyroid $P(S_{ms}|D_{hypo})P(D_{hypo}) =$

$(.001)(.15) = .00015 = .00076$

Euthyroid $P(S_{ms}|D_{euth})P(D_{euth}) =$

$(.107)(.65) = .06955 = .35000$

Hyperthyroid $P(S_{ms}|D_{hype})P(D_{hype}) =$

$(.648)(.20) = \underline{.12940} = \underline{.65000}$

Total .19910 1.00000

Thus a randomly selected patient with moist skin is more likely to be hyperthyroid than hypothyroid or euthyroid.

Using the information given above (patient's symptom is moist skin), we can also determine the probability with which a given patient found to have a particular symptom pattern will belong to each diagnostic group using Bayes' Theorem. An example follows:

$$P(D_r|S_p) = \frac{P(S_p|D_r)P(D_r)}{\Sigma[P(S_p|D_i)P(D_i)]}$$

$$P(D_{hypo}|S_{ms}) = \frac{P(S_{ms}|D_{hypo})P(D_{hypo})}{P(S_{ms}|D_{hypo})P(D_{hypo}) + P(S_{ms}|D_{hype})P(D_{hype}) + P(S_{ms}|D_{euth})P(D_{euth})}$$

$$P(D_{hypo}|S_{ms}) = \frac{(.001)(.15)}{(.001)(.15) + (.107)(.65) + (.648)(.20)} =$$

$$\frac{.00015}{.19910} = P(D_{hypo}|S_{ms}) = .0008$$

$$P(D_{euth}|S_{ms}) = \frac{(.107)(.65)}{(.19910)} = 0.3492$$

$$P(D_{euth}|S_{ms}) = \frac{(.648)(.20)}{(.19910)} = 0.6500$$

Thus, the patient will probably be hyperthyroid.[66]

In a later study, Overall, Williams and Fitzgerald applied the techniques presented above to a series of 210 new cases of suspected

---

[66]John E. Overall and Clyde M. Williams, "Conditional Probability Program for Diagnosis of Thyroid Functions," Journal of the American Medical Association, CLXXXIII (February 2, 1963), 309-11.

thyroid disease, using an 879 case data base on a computer. Of the 210 cases used, 11 errors were found due to systematic errors, four of which were classified as masked hyperthyroidism. A major revision was made, a print-out of which can be seen in Table XIII.[67]

TABLE XIII

Print-Out of Diagnosis of Thyroid Functional Status and Differential Diagnosis

Items		Percent Probability	Items Present
Signs and Symptoms	Hypothyroid	1	17
	Euthyroid	99	
	Hyperthyroid	0	
Laboratory Tests	Hypothyroid	5	3
	Euthyroid	95	
	Hyperthyroid	0	
Total	Hypothyroid	0	
	Euthyroid	100	
	Hyperthyroid	0	
No thyroid disease			100
Simple goiter (250)			0
Nontoxic nodular goiter (251)			0
Thyrotoxicosis (252)			0
Myxedema and cretinism (253)			0
Other diseases (thyroiditis) (254)			0
Carcinoma of thyroid (194)			0

In the new program the computer finds the probability of hyperthyroidism, euthyroidism, or hypothyroidism, first on the basis of the signs and symptoms alone, second on the basis of laboratory

---

[67]Williams, Fitzgerald, and Overall, Recent Advances, p.259.

values alone, and third on the basis of all available information. The separation of the three sets of data allows mistakes in diagnosis to be easily traced. It also allows possible differences in diagnosis due to laboratory tests, and signs and symptoms to be clearly seen. Another revision in the program was to enter the presence or absence of each symptom into the data matrix (data base) after the computer diagnosis of each case. This was an attempt to avoid the laborious job of retreiving charts and data sheets for each new type of sign and symptom and permits the rapid accumulation of exact probability estimates in a data matrix (data base) currently having a small number of entries. A program for the withdrawal of the sign and symptom entries when a mistake has been made was also written.[68]

Another study using the revised program (referred to as the accumulating/non-accumulating program) was applied to 500 cases of suspected thyroid disease at the University of Florida (including the 210 cases used previously in the trial run). Criteria of agreement between the computer diagnosis and the physician's diagnosis was that if the patient improved as a result of appropriate treatment, it was considered that the patient had the disease originally diagnosed by the computer. A total of 14 cases were misdiagnosed for an accuracy of 97 percent.[69]

One of the major values of the program developed by Fitzgerald, Williams, and Overall lies in the fact that rare disease will be

---

[68] Ibid., p. 259.

[69] Ibid., pp. 258-59.

considered and looked for automatically in every case considered in the future program. They point out that programmers of modest experience should have no difficulty in adding their own rare cases to the existing program.[70]

B. University of Texas Medical School Studies

While in Germany, Reichertz, Winkler and Kloss used the method developed by Fitzgerald, Williams and Overall to examine the efficiency of the program to compare the findings of two laboratories with different diagnostic strategy and possibly different base rates and to verify one of the most important facts resulting from application of computers, that is, the possibility of transferring experience of the research group to the other in the form of programs and basic data.

When 100 hyperthyroid cases were examined under the original program, 93 percent accuracy was achieved. When using the accumulating program, two previously incorrectly classified cases were given the correct diagnosis, thus increasing the efficiency to 95 percent.

Another test involving 105 non-selected consecutive patients of all three diagnostic categories was then processed without using laboratory values. Correct diagnosis was achieved in 91 percent of of these cases, which is above the diagnostic possibility (85-90

---

[70]Lawrence T. Fitzgerald, John E. Overall, and Clyde M. Williams, "A Computer Program for Diagnosis of Thyroid Disease," The American Journal of Roentgenology, Radium Therapy and Nuclear Medicine, XCVII (August, 1966), 904.

percent correct diagnosis) of a well-trained staff using 29 different signs and symptoms.[71]

When the 100 cases were accumulated there were many increases and a few decreases in the accumulated matrix of incidence data for hyperthyroid clinical signs and symptoms over the 879 original cases studied by Fitzgerald, Williams and Overall, as illustrated in Table XIV.[72] Reichertz, Winkler and Kloss point out that this indicates that the symptoms could have been misinterpreted by the resident who questioned the patients. In some instances the same patient questioned by two different interns gave a considerable number of divergent answers, "even when apparently unmistakable symptoms such as diarrhea or weight gain were to be checked." This resulted in several test formulations in which questions had to be asked word-by-word and hopefully allowed the least erroneous results. The main results of the study were summarized in Reichertz's, Winkler's and Kloss's paper and are presented in the authors' words.

1. The program lead to a convincingly high number of correct diagnoses in spite of the fact that the two laboratories used a different diagnostic strategy and that there could have been differences in laboratory ranges as well as in base rates.

---

[71] P. L. Reichertz, C. Winkler, and G. Kloss, "Experiments with a Computer for the Diagnosis of Thyroid Disease," Conference on the Use of Computers in Radiology (Chicago, October, 1966), E-62-E69.

[72] Ibid., p. E-67.

2. During our studies the incidence matrix was only slightly changed, so it can be assumed that the original definition of the different diseases was consistent with the frequency distribution of the different signs and symptoms in out patients.
3. Thus by the aid of a computer program the experience and efforts of two different research groups can be integrated. This is a minute but real step toward a higher level of complexity.[73]

TABLE XIV

Incidence Data for Hyperthyroidism, Clinical Signs and Symptoms

	Results by Fitzgerald and Williams	Own experiments added
Nervousness	0.915	9.925
Heat sensitivity	9.742	0.746
Perspiration	0.678	0.747
Appetite gain	0.605	0.579
Weight loss	0.836	0.846
Hyperkinetic movements	0.755	0.688
Warm, moist skin	0.708	0.760
Light finger tremor	0.871	0.898
Lethargy	0.001	0.001
Cold sensitivity	0.051	0.094
Decreased perspiration	0.001	0.001
Appetite loss	0.133	0.160
Weight gain	0.023	0.032
Slower movements	0.018	0.012
Dry, rough skin	0.009	0.011
Face edema	0.008	0.026
Eye symptoms	0.300	0.446

C. Cambridge Hospital Studies

Measurements of the free thyroxin, T_F, has been considered one of the most useful reflections of the clinical thyrometabolic

[73] Ibid., p. E-69.

states. An accurate estimate of the circulating T_F is possible by the competitive binding analysis (T_4(test and the resin uptake test, and therefore make the exact measurement of T_F unnecessary under normal clinical circumstances.[74]

Robin, Refetoff and Selenkow have used these tests in their research and have extended their methods to the use of a computer. The amount of free thyroxis, T_F, is determined from a computer-calculator after approximation tests are made. Values for the amount of T_F were found by multiplying the value of serum thyroxine concentration (T_4D) times the ratio of a standard reference serum (RT_4). These values (T_F) have been placed on a frequency distribution (Figure 11) which separates hyperthyroid cases, euthyroid cases and hypothyroid cases.[75] The dark area between hypothyroid and normal and between normal and hyperthyroid are borderline cases. This distribution was made by considering serum samples from 137 normal cases (no medication for the two months prior to the collection of the samples), 70 hyperthyroid, 50 hypothyroid, and 160 "normal" (euthyroid hospitalized or ambulatory patients; pregnant women, possible on medications of estrogen and androgens) subjects. A breakdown of the distribution from Figure 11 is shown in Table XV.[76] This table also shows the computer diagnosis for each of the ranges of the distribution.

---

[74]Noel L. Robin, Samuel Refetoff, and Herbert A. Selenkow, "The Computer in the Diagnosis of Thyroid Disease," Journal of Nuclear Medicine, XII (September, 1971), 622.

[75]Ibid.

[76]Ibid., p. 623.

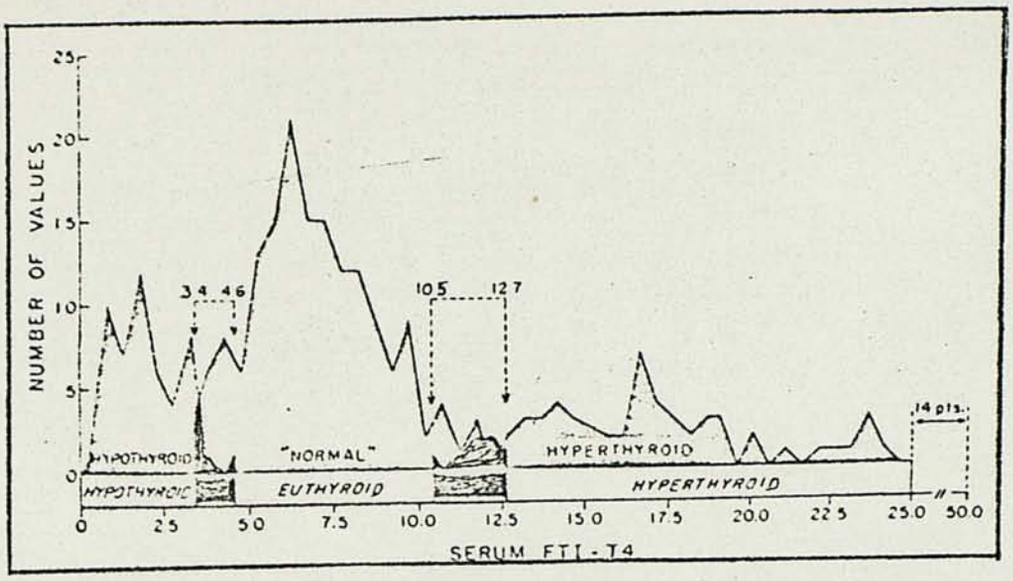

Figure 11. Frequency distribution of FT1-T, values of hypothyroid, "normal," and hyperthyroid patients in this study.

TABLE XV

Ranges of FT1-T_4 for Diagnostic Discrimination of Thyroid Function by Computer Analysis

FTI-T_4 range	Hypothyroid*	"Normal"*	Hyperthyroid*	Computer diagnosis
< 3.43	86.0% (43/50)	0.0% (0/160)	0.0% (0/70)	Hypothyroid
3.43– 3.84	12.0% (6/50)	2.5% (4/160)	0.0% (0/70)	Borderline hypothyroid-euthyroid
3.84– 4.56	2.0% (1/50)	6.9% (11/160)	0.0% (0/70)	Probable euthyroid
4.56–10.45	0.0% (0/50)	87.5% (140/160)	0.0% (0/70)	Euthyroid
10.45–12.73	0.0% (0/50)	3.1% (5/160)	7.1% (5/70)	Borderline hyperthyroid-euthyroid
>12.73	0.0% (0/50)	0.0% (0/160)	92.9% (65/70)	Hyperthyroid

* Figures in parentheses represent the number of patients in that category over the total number of patients in the entire diagnostic group.

D. University of Arkansas Medical Center Studies

Most of the numerical tests explained have been restricted to "yes-no" results. Oddie points out that most test results are distributed continuously and that a great deal is known about their

mean values and standard deviations.[77] Therefore he felt it might be worthwhile to develop a program based on this knowledge. He feels that his program supplements but does not duplicate methods described by others who have based the diagnosis on the presence or absence of clinical signs and symptoms, and upon "yes-no" rather than continuously distributed test results.

His method involves two basic assumptions.

> The first is that thyroid patients can be sorted into three different populations (hypothyroid, euthyroid, and hyperthyroid) having different test characteristics as was done in the previous tests. The second assumption is that the results of different thyroid function tests are independent which enables Bayes' theorem to be applied. This method gives full weight to tests of independent types, but those of similar nature are combined and considered as one test for processing.

Tests grouped in this way are shown in Table XVI.[78]

There are several other features of Oddie's program and these are listed below.

a) Diagnosis is suspended and messages are printed if data is missed or out of chronological order.
b) If the input comprises the initial measurements for radioiodine uptake tests, the observations are processed, the computed results are punched on a card for the data file, and a test report is printed with the diagnostic report.
c) When finished observations only are used as input, the diagnostic report is printed.
d) A separate diagnosis is reported whenever tests are spaced more than 6 weeks apart, and whenever the patient is treated by 131_I therapy or thyroidectomy.

---

[77] T. H. Oddie, "Computer Diagnosis from Tests of Thyroid Function," Journal of Clinical Endocrinology and Metabolism, XXXII (February, 1971), 168.

[78] Ibid.

TABLE XVI

Categories of tests

Group	Tests included	Limitations incorporated in program logic
1	Early clearance k_1	Very low values excluded if 24-hr uptake available
	24-hr uptake	Very high values excluded if k_1 available
	5-15 min clearance	Used if binding blocking drugs vitiate later tests
2	PBI BEI T_4-by column	Given equal weight unless differentially affected by medication. Limits adjusted for estrogens
3	T_3-resin uptake ratio	Limits adjusted for estrogens when necessary
4	T_4-by resin sponge	Temporarily separated from group 2 for purpose of study
5	TSH stimulation	Diagnostic only if minimal stimulation is found Excluded from hyperthyroid status Excluded if the patient had prior 131_I therapy or thyroidectomy
6	T_3 suppression	Excluded from low uptake subjects Excluded from hypothyroid status
7	T_4 suppression	As for T_3 suppression
	Achilles reflex	Not included in present program
	PBI-131_I	Not included in present program

e) Individual tests are ignored if coded as affected by medication, unless the diagnostic values can be adjusted to allow for its effect, or unless the results are applicable to stimulation or suppression calculations.
f) Comments are printed about the results that conflict with the overall diagnosis.
g) A diagnosis of hypothyroidism is not accepted for an untreated patient with a low radioiodine uptake without confirmation by other tests.
h) A diagnosis of hyperthyroidism is not accepted if a high PBI or similar test result is the only valid information available.
i) Diagnosis is withheld if the uptake is low and another tests, such as the protein-bound iodine (PBI), is high.
j) Diagnostic results are reported as below. In the doubtful cases the expected percentage of mid-diagnosed results is also printed, or further action is suggested.

Hyperthyroid

Doubtful high. Most probable result hyperthyroid but more than 1 percent of euthyroid patients might be mis-diagnosed or most probable result euthyroid but more than 5 percent of hyperthyroid patients might be mis-diagnosed.

Euthyroid

Doubtful low. Similar to doubtful high but 5 percent mis-diagnosis allowed in each bordering category.

Hypothyroid

Figure 12 shows an example of a combined test and diagnostic report.[79]

Results of the program when applied to 5010 cases studied at the University of Arkansas Medical Center since 1954 are shown in Table XVII.[80] Eliminated from the 5010 cases are 159 cases because no retrospective decisions could be made, 353 due to insufficient data

---

[79] Ibid., p. 168-70.

[80] Ibid., p. 170.

UNIVERSITY OF ARKANSAS MEDICAL CENTER
DIVISION OF NUCLEAR MEDICINE
THYROID PATIENT REPORT

UNIT NUMBER 323725

SEX = FEMALE RACE = NEGRO AGE = 58
DATE OF LAST TEST 4-18-1970
GOITER STATUS = NONE
RECORDED THYROIDAL STATUS = UNKNOWN

LATEST PROCEDURES OR TEST RESULTS

DATE	TREATMENT	TEST RESULTS
4-18-1970		24-HR ^{131}I UPTAKE = 43.5%
4-18-1970		T_4 BY RESIN SPONGE = 14.5 μg/100 ml

COMPUTER DIAGNOSIS BASED ON TECHNICAL TESTS

TEST TYPE	PROBABILITY OF STATUS EUTHYROID	HYPOTHYROID	HYPERTHYROID
UPTAKE	0.693	0.000	0.307
T_4 SPONGE	0.696	0.000	0.304
ALL TESTS	0.569	0.000	0.431

DIAGNOSIS FROM COMBINED TESTS IS BORDERLINE HIGH
WE RECOMMEND A SUPPRESSION TEST FOR MORE RELIABLE DIAGNOSIS

Figure 12. Example of combined test and diagnostic report.

Table XVII

Comparison of Computer Diagnoses With Accepted Diagnoses

Accepted diagnosis	Classification by computer: Hypothyroid	Borderline low	Euthyroid	Borderline high	Hyperthyroid	No diagnosis
Hypothyroid	183	11	57	1	0	42
	7	0	0	0	0	0
Euthyroid	23	26	3251	212	48	236
	1	1	127	2	0	3
Hyperthyroid	1	0	56	104	684	75
	0	0	1	1	43	9
Undecided	0	8	50	5	6	90
	0	0	2	0	1	2

New series data with undecided results included and borderline results pooled

Accepted diagnosis*	Classification by computer: Hypothyroid	Euthyroid	Hyperthyroid
Hypothyroid	7	0	0
Euthyroid	2	127	2
Hyperthyroid	0	1	44

for diagnosis, and 353 for borderline diagnoses. Oddie points out that presumably all of these could have been classified definitely by using more extended tests. Out of the remaining diagnoses 186 were wrong and 4118 correct for a figure of 82 percent correctly diagnosed. If sufficient test data were supplied for all null and borderline diagnoses to be eliminated, the percent correct would increase by almost 14, bringing the total to 95.7 percent. Similar results were obtained in a prospective test of the system with data for 200 new patients not used in developing the program, which can also be seen in Table XVII as the second line in each category.[81]

Additional studies were made after minor changes in the basic program. The first study involved diagnostic success using only a single test, which was allowed by "removing the restrictions about not diagnosing from high PBI values or low 131_I uptakes." Table XVIII shows the T_3 resin sponge uptake as the worst of the groups and the T_4 resin sponge test groups as the best.[82]

The second study involved diagnosis from a selected set of tests. It has already been implied that some tests could be omitted without much loss of diagnostic accuracy.

> To study this possibility, Oddie retained the 131_I test because it is used not only alone, but also as the first step in secondary tests such as TSH stimulation, T_3 and T_4 suppression, and $KCLO_4$ discharge measures. The T_4 sponge test was also retained, being the best. However, the PBI results were deleted if a T_4 sponge result was available,

---

[81] Ibid.

[82] Ibid.

and the T_3 resin results were deleted if either a T_4 sponge or a PBI result was available.[83]

The diagnostic results of this test were slightly worse than that for the same 200 previous cases, increasing from 1 to 1.5 percent in the wrong category and an increase by 2 percent in the borderline category.

TABLE XVIII

Comparison of Unrestricted Diagnostic Results of Single Tests

Test group	No. of results analyzed	Percent of diagnoses		
		Borderline	Correct	Wrong
1. 131_I uptake	1316	22.4	72.0	5.6
2. PBI (Observer 1)	383	14.9	82.2	2.9
2. PBI (Observer 3)	728	17.4	77.5	5.1
3. T_3 resin uptake	113	35.4	52.2	12.4
4. T_4 resin sponge	189	7.4	86.2	6.4
5. TSH stimulation	104	25.0	63.5	11.5
6 & 7. T_3 and T_4 suppression	93	15.1	76.3	8.6
For comparison: computer with all tests available (new series)	183	2.2	96.7	1.1

The third and final test was concerned with trends in test values which could be detected in different groups and also with changes in time. "Radioiodine uptake values changed in Arkansas from 1960 to 1963, because the mean dietary intake of iodine increased." Oddie points out that test values depend significantly upon the patient's race and the county within the state in which he resides.

---

[83] Ibid., p. 171.

The information supplied by Oddie is only a beginning, for he has stated that there are plans to apply the system described to a larger series of new patients and to supplement the best diagnosis with a computer-calculated clinical diagnosis closely following the method of Overall and Williams.[84]

6. Computer Diagnosis in Dermatology

Arthur L. Norins of the Indiana University Medical Center conducted a pilot study to examine how a computer might be used for diagnosis in dermatology. "The study involved a delineation of the important characteristics of disease states, a method of storing data, and a means of comparing patient data with disease data in order to make a diagnosis."

Norins defined the word "attribute" in his study to denote a general term for any information concerning a patient, regardless of its source (from history or examination). In order to keep the study simple, a fixed number of attributes were chosen. This places great importance on the attributes chosen and therefore allows the possibility of including unimportant attributes and excluding important ones. To combat this possibility, "The system developed here is subject to adjustment whereby a seemingly important attribute can be dropped if it is later felt to be unimportant, and there is a provision for addition of new attributes to the list."

The example shown in Figure 13 is a partial list of attributes which was developed by a physician's historical information and from a

---

[84] Ibid., pp. 167-71.

physical examination of the patient.[85] The list of clinical attributes, along with a histological attribute list was evaluated for several hundred diseases and placed in a "catalog" of disease attributes. The cataloging of the information was the first step in the pilot study.

Clinical attribute sheet #1

Number	Attribute
	RACE INCIDENCE
1	Negro
2	White
3	Other
	SEX INCIDENCE
4	Male
5	Female
	AGE AT ONSET OF CONDITION
6	Less than one year
7	Less than 20 yr. more than 1 yr.
8	Less than 65 yr. more than 20 yr.
9	Over 65 yr.
	DURATION OF CONDITION
10	Less than 3 months
11	Longer than 3 months
	SYMPTOMS
12	Pain (Mild to Moderate)
13	Pain (Severe)
14	Pruritus (Mild to Moderate)
15	Pruritus (Severe)

Figure 13. Partial list of clinical attributes.

[85]Arthur L. Norins, "Computers in Dermatology," Archives of Dermatology, XC (November, 1964), 507.

The next step in the study was the evaluation of the attributes schedules for individual patients (Table XIX) and for individual diseases (Table XX).[86]

TABLE XIX

Evaluation of Several Attributes for a Particular Patient

Attribute	(+1) Yes	(-.5) No	(0) Unknown	Score
A	X			1
B		X		- 0.5
C			X	0
D		X		- 0.5
E	X			1

TABLE XX

Evaluation of Several Attributes as Seen in a Specific Disease

Attribute	0% 10% (-3)	10% 30% (-1)	30% 70% (0)	70% 90% (+1)	90% 100% (+3)	Score
A					X	3
B				X		1
C		X				- 1
D	X					- 3
E			X			0

In the evaluation process, each attribute was given a positive value if present and a negative numerical value if not present. In this example, the patients attributes if present were given a value of +1 if present, -.5 if not present and 0 if unknown. The patient value and

---

[86] Ibid., p. 508.

disease value of an attribute are multiplied together (Table XXI).[87] Their product will be positive if the signs of the numerical values for the patient and disease are the same. The product will be negative if the signs of the numerical values are different. The score column is then added up. The larger the positive value, the greater the similarity. The ratio between the patient's score for a disease and the perfect score for a disease (to be found in the disease attribute catalog) is determined with the largest ratio considered the most likely exhibited by the patient.[88]

TABLE XXI

Comparison of the Information From Specific Disease With the Information From a Particular Patient

Attribute	Disease	Patient	Score
A	3	1	3
B	1	- 0.5	- 0.5
C	- 1	0	0
D	- 3	- 0.5	1.5
E	0	1	0
			4.0

Table XXI shows that the total score for the patient and disease when added up is 4.0. This figure can be compared to the perfect score of the disease being considered to determine the ratio mentioned above. Although the perfect disease scores were not given, Table XXII, shows the ratio determined for three diseases.[89] The disease with the highest ratio is chosen as the disease most likely

---

[87] Ibid.

[88] Ibid., pp. 509-10.

[89] Ibid., p. 510.

possessed by considering the patient's attributes, which in this case is disease number 135 with a ratio of 0.9. Since it is necessary to compare a patient, attribute by attribute, to each possible disease as was done above, the possible combinations which must be compared is very large and therefore should show why a computer is involved.

TABLE XXII

The Final Score Calculated for Several Different Diseases for the Patient Being Evaluated

Disease Number	Ratio
135	0.9
24	0.7
73	0.6

This program was tested on twenty-five dermatology cases of which eighteen, or 72 percent, were diagnosed correctly. Four of the misdiagnosed cases were diseases not included in the catalog of attributes developed.[90]

7. Computer Diagnosis of Primary Bone Tumors

Lodwick, Haun, Smith, Keller and Robertson have used the logic of Bayes' Theorem to correlate eight possible diagnoses to seventy-one units of clinical and radiological data. Although the authors do not present the actual data used in their study, they do present a summary of seventy-seven cases tested and give a hypothetical sample to illustrate the application of their method.

---

[90] Ibid., p. 506-11.

In the hypothetical case, there are three possible diagnoses - Y_1, Y_2, and Y_3 - and nine possible units of clinical or radiological data - X_1 to X_9. The possible diagnoses and the clinical and radiological data are presented in a probability matrix as shown in Figure 14.[91]

	Incidence	Age		Location		Size		Matrix	
		01-20	21-40	Epiphysis	Metaphysis	0-6.0	6.1-12	Recalcification	Flocculent
	X1	X2	X3	X4	X5	X6	X7	X8	X9
Y1 Giant Cell Tumor	30	15	85	90	95	04	85	99	01
Y2 Codman's Tumor	10	85	15	98	35	95	05	65	35
Y3 Chondrosarcoma	60	30	70	01	99	20	80	35	65

Figure 14. Hypothetical probability matrix.

Hypothetical probability values reflecting the relative frequency with which each attribute is found in each diagnosis are placed in the individual squares. Thus in the hypothetical probability matrix (Figure 14), 85 percent of the Giant Cell tumors are found in persons between the ages of twenty-one and forty, compared with 15 percent of Codman's tumors and 70 percent of chondrosarcoma. These probabilities given in Figure 14 can be used together with Bayes' Theorem to calculate the probabilities that the patient possesses a given disease.

---

[91]G. S. Lodwick, et al., "Computer Diagnosis of Primary Bone Tumors: A Preliminary Report," Radiology, LXXX (February, 1963), 274.

The hypothetical case is now presented to demonstrate the technique used in the computer diagnosis of primary bone tumors. A twenty year old patient (X_2) having a bone tumor involving the epiphysis (X_4) and the metaphysis (X_5) 6 cm. in diameter (X_6), and demonstrating a floucculent pattern (X_9), can be diagnosed using Bayes' Theorem. The probability for each of the three possible diseases is calculated below:

Bayes' Theorem adapted to diagnosis is:

$$P(\text{disease}_r|\text{attributes}_{1\ldots i}) =$$

$$\frac{P(\text{disease}_r) \cdot P(\text{attribute}_1|\text{disease}_r) \cdot \ldots, \cdot P(\text{attribute}_i|\text{disease}_r)}{\sum_{\substack{\text{all } i \\ \text{all } n}} [P(\text{disease}_n) \cdot P(\text{attribute}_1|\text{disease}_n) \cdot \ldots, \cdot P(\text{attribute}_i|\text{disease}_n)]}$$

The probability values $P(\text{attribute}|\text{disease}_i)$ for attributes X_2, X_4, X_5, X_6 and X_9, which are present in the patient being considered, can be found in Figure 14 $P(X_2|Y_1) = 0.15$. The values for $P(Y_1)$ are found in column X_1 for each of the three rows, Y_1, Y_2, Y_3, respectively $P(Y_1) = 0.30$.

$$P(Y_1 \mid X_2,X_4,X_5,X_6,X_9) =$$

$$\frac{P(X_2|Y_1) \cdot P(X_4|Y_1) \cdot P(X_5|Y_1) \cdot P(X_6|Y_1) \cdot P(X_9|Y_1) \cdot P(Y_1)}{\begin{aligned} &P(X_2|Y_1) \cdot P(X_4|Y_1) \cdot P(X_5|Y_1) \cdot P(X_6|Y_1) \cdot P(X_9|Y_1) \cdot P(Y_1) \\ &+ P(X_2|Y_2) \cdot P(X_4|Y_2) \cdot P(X_5|Y_2) \cdot P(X_6|Y_2) \cdot P(X_9|Y_2) \cdot P(Y_2) \\ &+ P(X_2|Y_3) \cdot P(X_4|Y_3) \cdot P(X_5|Y_3) \cdot P(X_6|Y_3) \cdot P(X_9|Y_3) \cdot P(Y_3) \end{aligned}}$$

$$P(Y_1 \mid X_2,X_4,X_5,X_6,X_9) = \frac{(0.15)(0.90)(0.95)(0.04)(0.01)(0.30)}{\begin{array}{l}(0.15)(0.90)(0.95)(0.04)(0.01)(0.30) \\ + (0.85)(0.98)(0.35)(0.95)(0.35)(0.10) \\ + (0.30)(0.01)(0.99)(0.20)(0.75)(0.60)\end{array}} =$$

$$P(Y_1|X_2,X_4,X_5,X_6,X_9) = \frac{0.0000153900}{0.009410875} = .0015484$$

$$P(Y_2|X_2,X_4,X_5,X_6,X_9) =$$

$$\frac{P(X_2|Y_2) \cdot P(X_4|Y_2) \cdot P(X_5|Y_2) \cdot P(X_6|Y_2) \cdot P(X_9|Y_1) \cdot P(Y_2)}{.009410875}$$

$$\frac{(0.85)(0.98)(0.35)(0.95)(0.35)(0.10)}{.009410875} = 0.975$$

$$P(Y_3|X_2,X_4,X_5,X_6,X_9) =$$

$$\frac{P(X_2|Y_3) \cdot P(X_4|Y_3) \cdot P(X_5|Y_3) \cdot P(X_6|Y_3) \cdot P(X_9|Y_3) \cdot P(Y_3)}{.009410875}$$

$$P(Y_3|X_2,X_4,X_5,X_6,X_9) = \frac{(0.30)(0.01)(0.99)(0.20)(0.65)(0.60)}{.009410875} =$$

$$= 0.023$$

Thus there is a .158 percent, 97.5 percent, and 2.3 percent probability that Y_1, Y_2, and Y_3 respectively, is the correct diagnosis.

Through use of a computer the method presented above was applied to seventy-seven cases with a probability matrix of sixe 8 by 71, that is, eight possible diagnoses and seventy-one findings. The seventy-one findings (symptoms) were not presented in the article due to lack of space. Table XXIII displays the eight possible diagnoses and a summary of the cases considered.[92] Seventy-seven point nine percent of the cases being considered were correctly diagnosed.

---

[92] Ibid., p. 275.

TABLE XXIII

Results of Computer Diagnosis of 77 Bone Tumors

	No. of Cases	No. of Cases Correctly Diagnosed	Percent of Cases Correctly Diagnosed
Chondroblastoma	8	7	87.5
Chondrosarcoma	11	7	63.6
Ewing's sarcoma	11	9	81.8
Fibrosarcoma	11	9	81.8
Giant-cell tumor	12	12	100.0
Osteosarcoma	14	11	78.6
Parosteal sarcoma	2	2	100.3
Reticulum-cell sarcoma	8	3	37.5
	77	60	77.9

The results of the study were considered encouraging by the authors, with the errors recognized partially due to "imperfect programming."[93]

[93] Ibid.

PART III

SURVEY ON MEDICAL DIAGNOSTIC SOFTWARE

A survey on medical diagnostic software titled "Computer Assisted Diagnostic Software" was sent to a sample of physicians in an attempt to define the status of diagnostic software as a tool for physicians in the real world. The questions were designed to determine physician's interest in using diagnostic software, their knowledge of the subject, their patient's predicted reaction, and its present use by physicians.

CHAPTER VI

SURVEY OF THE ORLANDO, FLORIDA, PHYSICIANS CONCERNING USE OF MEDICAL DIAGNOSTIC SOFTWARE

One hundred physicians were selected randomly from the December 1971-1972 Orlando, Florida, Telephone Directory. The physicians were listed in alphabetical order under the section titled "Physicians and Surgeons - MD" and were numbered from 1 to 356, start-ine with the first physician. A random digits table was used to select 100 of the 356 physicians.

Forty percent of the questionnaires were returned. The results of this survey are as follows. (A copy of the survey and cover letter can be found in Appendix A.)

1. Over 89.7% of the physicians in the area surveyed have never used a computer diagnostic program to assist them in diagnosis. 7.7% have used a computer diagnostic program to assist in diagnosis and 2.6% did not comment on the question.

2. Forty-three point six percent of the physicians knew of a diagnostic program presently being used in the area surveyed. 48.7% did not know of a program being used and 7.7% did not comment.

3. Sixty-one point five percent of the physicians showed that they would be willing to use a computer to assist them in diagnosis if economically feasible. 33.3% showed that they would not be willing and 5.1% did not comment.

4. Seven point seven percent of the physicians felt that the diagnostic programs presently being used could be revised to be of more use by a wider application, 17.9% felt by better accuracy, 35.9% felt by greater accessibility, and 5.1% felt by having it simpler to use. The comments listed in the blank under "others" were: increase efficiency; no way; personal interviews; apply to cardiology; use as physician assistant; and stop investigation.

5. About 71.8% of the physicians felt that the program should aim its use at physicians, 17.9% felt the program should aim its use at nurses, and 23.1% felt that the program should aim its use at technicians. Several of the physicians checked more than one of the items which will cause the total percentage to be greater than 100%. 17.9% of the physicians did not comment at all on the question.

6. The answers to this question could easily be classified into the following three categories: positive, negative, and unsure. 48.7% of the physicians felt their patients would accept a computer to assist in diagnosis, 28.2% felt their patients would reject its use, and 23.1% were unsure about their patients feelings.

7. The following comments were made in response to question number 7 which asked for advice on how to help make a diagnostic computer program a useful aid to physicians.

- offer program to area physicians as a major aid in diagnosis; relate to hospital if possible;
- more information;
- consult with a variety of specialists to determine most effective means of providing this assistance;
- it should be a useful aid and nothing else;
- I am a person who believes in personal contact with patients to evaluate the patient.
- I'm terribly sorry - I don't know enough about it to be of help.
- not able to advise;
- make it economical, accurate, and report fully to patient's physician;
- practical;
- do case studies;
- small inexpensive and accurate;
- computer would make my practice too impersonal'
- reduce cost;
- more publicity;
- quick results, inexpensive and simple;
- assist in history taking information;
- good programming;
- make it easy and economical;
- stick to basics and do not try to get a machine that gives you a diagnosis;
- there is no way;
- in our field, diagnosis is not as difficult as management and treatment is more of a problem;
- diagnosis is not the problem - paper work - larger insurance forms - more government restrictions and red tape are the cause of physician shortage;
- most effective as a reliable history bank.

DISCUSSION OF RESULTS AND CONCLUSIONS

The studies which have been discussed and the results of the survey conducted have been useful in presenting the present role of computer diagnosis in the medical profession. The cross-section of literature allowed a close look at some of the more advanced studies which have been conducted and the results of the survey gave some insight into the physician's feeling concerning the real-life role medical diagnostic software presently plays. A closer look will now be taken and certain conclusions drawn.

The section of literature concerning multiphasic health testing shows one of the most successful areas which has been attempted in real-life applications of medical diagnostic software. In 1970, 101 multiphasic health testing programs were in operation in the United States with 21 more planned for operation by 1972. The only other actual applications of medical diagnostic software presented in this paper are the thyroid program and the acid-base disorder program which is set up on a time sharing computer terminal. The two methods are being primarily used in the medical schools where their application was developed. The other software methods discussed in this paper are still undergoing research. Medical Data Screening has been used in research studies but no real-life application was discussed by the author. H. C. Becker and Allen L. Hammond in their articles on computer diagnosis in heart disease emphasized the preliminary nature of their studies and expressed confidence that with the future studies

planned, this method of computer diagnosis would be successful in their area of interest as well as other areas of diagnosis. Robert S. Ledley presented his paper on chromosome analysis, stating that more case studies would have to be made in order to prove its reliability. The authors of the pelvic surgery software, Neurath, Enslein, and Mitchell felt that their method was "good enough to establish a working system" and that a successful implementation of their system in pelvis surgery would pave the way for its adoption by other areas of diagnosis. Arthur L. Norris communicated that there are no further studies planned to follow up his work in applying computer diagnosis to dermatology. The authors of the bone tumor study suggested that the scope of future studies be enlarged to include behavior predictions of tumors, but did not state future plans concerning other studies. The reliability and efficiency of each software application must yet be proven, but the main portion of the work, writing the software and conducting pilot studies, has already been accomplished.

There are several basic logic methods which have been used in medical diagnostic software, but the statistical theory using Bayesian probability has been the most widely used. Although there is question about its validity due to the assumption concerning the independence of the variables (assumed that symptoms are independent of each other), it is still presently being referred to by many researchers as the most superior technique which has been found and therefore, will probably continue to be used.

The survey which was sent to 100 physicians in the Orlando, Florida, area was responded to by 40 physicians. This response, 40%

is a little above the average response for surveys. This implies that the physicians are interested in the improvement of the medical profession and therefore responded with a serious attitude which tends to amplify the validity of the answers received.

The introduction section of this paper pointed out that medical diagnostic software is in its stage of genesis; that possibly the topic has not had enough time, nor has enough research been conducted to reach a level of sophistication and reliability. The attitude of the authors of the research tends to reiterate this thought in that the studies have been presented in an educational or research tone. It seems as though the authors of the studies realize that this method of diagnosis is still to be considered a "Research Topic" and are displaying the accomplishment which has been achieved.

The lack of maturity which was suggested in the introduction is evident in the survey in which 89.7% of the physicians who returned their survey had never used a computer diagnostic program to assist them in diagnosis. Over 48% of the physicians were not aware that this method was being used in the Orlando area, while multiphasic health testing facilities were in operation in that area almost six months prior to the existence of the survey.

More time is needed to perfect the software, and therefore, more studies will have to be conducted in order to produce an efficient diagnostic tool. A feeling of confidence by physicians must be achieved by presenting successful studies which can be critiqued and proven to meet a high level of reliability so mistakes and misdiagnosis are few and far between.