AN AUTOMATED TECHNIQUE

FOR PATIENT HEALTH ANALYSIS

By

BERNARD J. SCHROER

Bachelor of Science in Engineering
Western Michigan University
Kalamazoo, Michigan
1964

Master of Science in Engineering
University of Alabama
Tuscaloosa, Alabama
1967

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1972

OKLAHOMA
STATE UNIVERSITY
LIBRARY

AUG 16 1973

AN AUTOMATED TECHNIQUE

FOR PATIENT HEALTH ANALYSIS

Thesis Approved:

James K. Eldin
Thesis Adviser

M. P. Terrell

James E. Stamford

Fred Blocks

Norman N. Durham
Dean of the Graduate College

861378

PREFACE

This research consists of the development of a computerized technique to assist the physician in analyzing a patient's health. Using the patient's physical examination as the primary source of input, this technique assists the physician in analyzing a patient's health by 1) summarizing the patient's prior medical data, 2) determining if the patient is normal or abnormal, 3) identifying those clinical variables which are significantly affecting the patient's health, and 4) observing any significant longitudinal drift in the patient's health.

I wish to acknowledge my sincere appreciation for the assistance provided by my advisory committee: Dr. Hamed K. Eldin, Committee Chairman and Thesis Adviser; Dr. James E. Shamblin and Dr. M. Palmer Terrell, of the School of Industrial Engineering and Management; Prof. Fredrick M. Black of the Computing and Information Sciences Department; and the late Dr. Wilson J. Bentley of the School of Industrial Engineering and Management.

The support of the George C. Marshall Space Flight Center (MSFC) is gratefully acknowledge, with special thanks

to James H. Spraul, M.D., Director of the MSFC Medical Center; and Mr. L. D. Martin of the MSFC Computational Laboratory.

The encouragement and suggestions of Wallace B. Frierson, M.D.; Louis B. Arnoldi, M.D., Director of Occupational Medicine and Environmental Health, NASA Headquarters; and Albert Oberman, M.D., Department of Public Health and Epidemiology, University of Alabama, is gratefully acknowledge.

I am indebted to Computer Sciences Corporation, and especially Mr. H. R. Wallace, for his personal assistance throughout my research.

Finally, and most importantly, my deep appreciation is extended to my wife, Kathleen, our daughter, Shannon, and our son, Bradley, for their understanding, encouragement, and many sacrifices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

FORMULATION OF THE PROBLEM

Origin of the Problem

The health of a nation's people has traditionally been measured by the rate at which they die. A nation's death rate which is decreasing is considered an indication of an increase in the nation's health. However, in recent years the concepts of health have undergone significant changes. No longer are decreasing death rates being considered as a measure of health. Instead, emphasis has shifted to the total quality of life rather than the length of life and to the positive elements of good health rather than merely the absence of disease.

It has been suggested that in the future it may be possible to compile a health index similar to the gross national product index. Such an index could be called the gross national health deficit. This index would combine all the days of healthful living which are lost each year by the sick, the days of life lost through death that comes too soon, and all the impairments suffered because of lack of

medical treatment and advice. Such an index still stresses, as do conventional health statistics, the negative aspects of health.

The applications of computers to medical research and practice are relatively new. In the early 1960's computers were almost unknown in medicine. Today the situation has changed considerably. Many leading hospitals use computers for many different purposes from calculating doses of medicines to planning menus. However, the use of computers to assist the physician is still in its infancy. There is little doubt that because of the rapid increases in the volumes of medical data, the physician of the future, and also of the present, will need new methods of arriving at diagnoses. Computers, with their large storage capacities and their fast computational capabilities, offer the means of assisting the physician.

Before computers can be used to assist the physician in diagnosing, it is important to understand the process used by physicians in making a diagnosis. An oversimplified explanation of the diagnosis process follows. First, the physician obtains the case facts from the patient's history questionnaire, physical examination, laboratory tests, etc. Second, he evaluates the relative importance of the different signs and symptoms. Some of the data may be more

heavily weighed than other data in his evaluation. Third, the physician makes a diagnosis. The diagnosis consists of listing all abnormalities which the case can resemble. Then, by an exclusion process on the compiled list of abnormalities, a specific abnormality is determined; or it may be that the abnormality cannot be determined.

Quite often the physician, after seeing the patient, has a "feeling" about the case. This "feeling" which is difficult to explain, is generally a summation of the physical impressions concerning the way the data seems to fit together, the patient's reliability, general appearance, facial expressions, and so forth.

Errors do occur in diagnosing; however, it is widely believed that the majority of errors result from excluding possible abnormalities during the diagnosis than from any other source. It is here where the computer can provide a valuable service to the physician by reminding him of all possible conditions and abnormalities associated with the case.

One area where computers have tremendous potential in assisting the physician is in analyzing a patient's health. Such an analysis of a patient's health is commonly accomplished through a routine physical examination. A phrase in vogue for this area of analysis is multiphasic

health screening. The concept behind multiphasic health screening is to detect any abnormalities that a patient may have with the minimum number of tests and for a nominal fee. Those patients having abnormalities, or possible abnormalities, are then referred for further and more detailed examination.

The computer's role in multiphasic health screening is illustrated by the flow diagram in Figure 1. A patient arrives at a medical clinic for a routine physical examination. Upon arrival, the patient completes a medical history questionnaire. The responses to the questionnaire are stored in the computer. A routine physical is then administered to the patient. Medical technologists generally administer the majority of the tests without the physician's assistance. A typical physical includes such tests as vision, hearing, physical characteristics, blood chemistry, urine analysis, chest x-ray, electrocardiogram, and pulmonary functions. The results of these tests are also stored in the computer.

After the physical examination the patient sees the physician. The physician reviews the medical history questionnaire and the results of the physical and then makes a diagnosis. To assist him in his diagnosis, the physician has access to the computer. Since all the patient's data

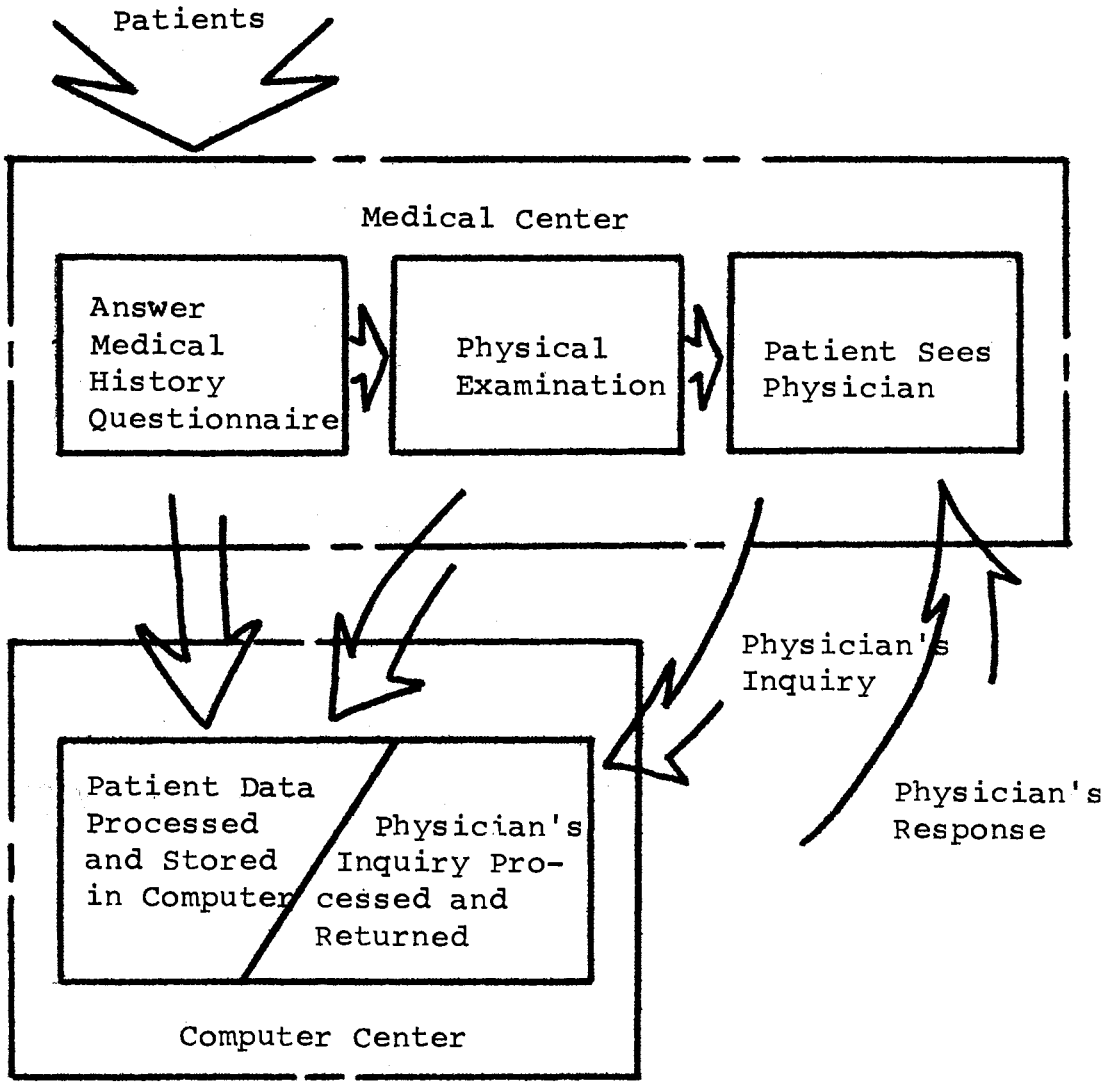


Figure 1. The Computer's Potential in Screening Programs

is stored in the computer, the computer is used as a tool to assist the physician in analyzing the patient's health.

For example, the patient's measurements could be compared with predefined norms, corresponding to his sex and age, and be displayed via a cathode ray tube. Those measurements falling outside the predefined limits could be flagged. Various statistical tests could also be made on the patient's data. Measurements from the patient's present physical could be compared and displayed with his previous physicals in hopes of detecting any drift in the patient's health. Also, significant data from the patient's past medical history could be retrieved from the computer and brought to the physician's attention.

An example of the multiphasic screening concept can be seen at the Kaiser Permanente Clinic in California (1) (2). This clinic is often considered the birthplace of multiphasic screening. At the clinic patients register at a rate of two every five minutes. Patients undergo a battery of medical tests, passing from one test station to another like parts in an assembly line. Computers are extensively used to store, retrieve, and analyze the patient's data. A similar program is being operated by the Tennessee Valley Authority (3). The TVA is using mobile multiphasic testing facilities.

Research Objective

Evidence of an abnormality in a patient presumes a knowledge of the same patient in a different state of health. This difference may be noted by observing the patient when he was normal. Or, this difference may be noted by comparing the patient to a hypothetical group who possess similar characteristics, such as sex and age.

A phenomenon of most physiological variables is that the distributions of these variables are generally normal. Therefore, it is possible to define a tolerance region which would contain a certain percentage of the values. Likewise, given a value of a clinical variable, a statistical test could be made to determine, with a given confidence, if the value is within defined limits.

One technique for detecting a patient's abnormalities is to compare each of his measurements against a distribution of that measurement for which the patient might be considered a member.

If only one variable is used to classify a patient as normal or abnormal, the patient's value could be compared with the distribution of that variable as in Figure 2. For any value falling within the tolerance region the patient would be classified as normal. Likewise, for any value

outside the tolerance region the patient would be classified as abnormal.

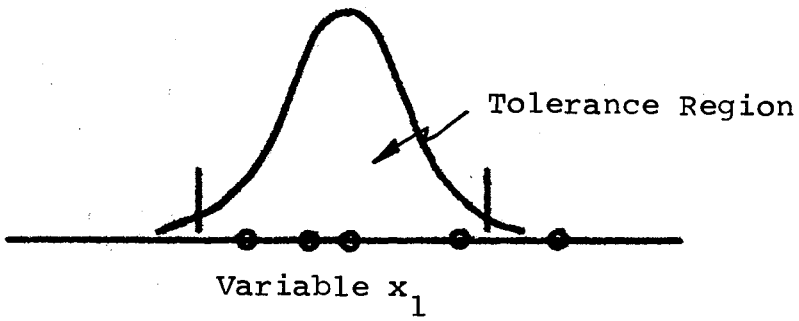


Figure 2. Single Variable Tolerance Region

If two clinical variables are used to classify a patient as normal or abnormal, the tolerance region would appear as a rectangle as shown in Figure 3. The patient's two variable observation can be plotted as a point in a plane. The assumption has been made that the two variables are independent. This is a common assumption in developing techniques for analyzing clinical variables. However, such an assumption does not consider the possible correlation which may exist between the variables.

If the dependence between the variables is considered, the tolerance region for the two variable observation is no longer a rectangle. Instead, the tolerance region becomes an ellipse as shown in Figure 4.

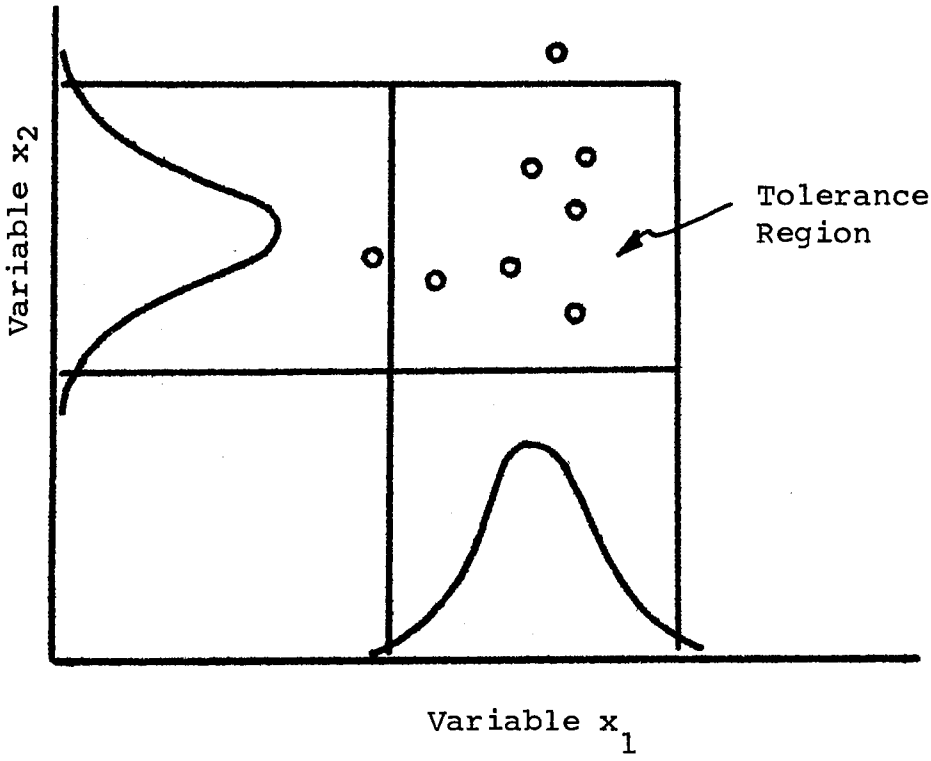


Figure 3. Two Variable Tolerance Region

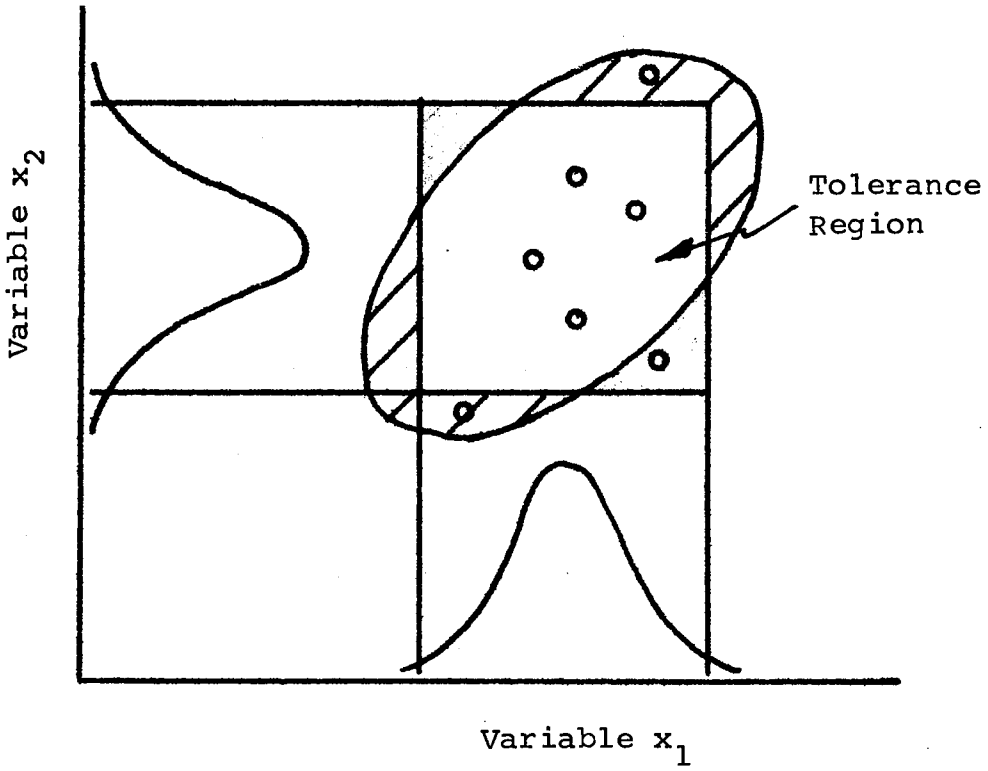


Figure 4. Bivariate Ellipsoidal Tolerance Region

A two variable observation falling in the cross hatched areas within the ellipse could be quite normal, but would be considered abnormal when compared with the tolerance limits in Figure 3. However, more seriously, a two variable observation falling in the shaded area outside the ellipse might be abnormal, but would be considered normal when compared with the limits of the individual variables.

In actual practice more than two variables are considered in determining if a patient is normal or abnormal. In general, an n-dimensional tolerance region must be considered. A patient's n-variable observation can then be tested to see if it falls within the tolerance region. If the point falls within the tolerance region, the patient would be classified as normal; otherwise, he would be classified as abnormal.

Once the patient has been classified as abnormal, a next logical step is to determine which of the clinical variables contributed most significantly to the patient's abnormal classification. By identifying those significant variables it should be possible to diagnose the patient's abnormality.

Another area of considerable interest is to obtain some measure of the patient's health and to observe this measure as a function of time. By observing any

longitudinal drift, it should be possible to detect a change in the patient's health.

In summary, the purpose of this research is to develop, for use primarily in occupational or multiphasic clinics, a computerized technique to assist the physician in analyzing a patient's health. Using the patient's physical examination as the primary source of input, the automated technique will assist the physician in analyzing a patient's health by 1) summarizing the patient's prior medical data, 2) determining if the patient is normal or requires additional medical attention, 3) identifying those clinical variables which are significantly affecting the patient's health, and 4) observing any significant longitudinal drift in the patient's health.

Related Research

Early interest in the subject of logical analysis of medical diagnosis may be attributed to the realization by some physicians that some sort of device was needed to aid diagnosis. The earliest diagnostic aids were handbooks listing the various signs and symptoms associated with various abnormalities.

In 1954 an English physician constructed a mechanical device similar to a slide rule which enabled a physician

to match various combinations of eighty-two signs and symptoms in order to choose the most likely diagnosis from a possible 337 diseases. Such a device could be helpful to a physician during consultation; however, it could hardly be considered an entirely satisfactory crutch for the physician's memory.

Other investigators in the 1950's devised mechanical methods using cards. Initially, card and needle systems were used. Later punched cards and card sorters were introduced. Note that all the above approaches did not utilize probabilistic relationships between symptoms and diseases. It was not until the introduction of the digital computer that probabilistic relationships could be considered.

Since the introduction of the digital computer in the early 1960's, a variety of approaches have been applied to the area of medical diagnosis. Several of these approaches are briefly discussed in the following paragraphs.

Probabilistic Approach

Warner (4) was the first investigator to successfully use a Bayesian conditional probability model, with the assistance of computers, to diagnose heart disease. Since 1961 the Bayesian approach has been applied to the

diagnosis of bone tumors (5), the classification of psychiatric patients (6), and the diagnosis of thyroid function (7).

The Bayesian approach does have its limitations. It requires that the diseases in question be mutually exclusive and that the symptoms be independently distributed given the diseases.

Using the Bayesian approach the problem is to determine the probability that the patient has disease D_j, when it is known that the patient has symptom S_i. The data upon which $P(D_j|S_i)$ is derived comes from medical knowledge. Such medical knowledge is generally given in the form of conditional probabilities: namely, the probability $P(S_i|D_j)$ that a patient having disease D_j, will have symptoms S_i.

Therefore, if medical knowledge is in the form of $P(S_i|D_j)$, the problem is to determine the diagnosis $P(D_j|S_i)$: namely, the probability of having disease D_j given that the patient has symptoms S_i. Using the Bayesian approach, $P(D_j|S_i)$ is computed as

$$P(D_j|S_i) = \frac{P(D_j)\ P(S_i|D_j)}{\sum P(D_j)\ P(S_i|D_j)}, \qquad (1)$$

where the summation is over all possible diseases under consideration.

Decision Table Approach

Decision tables have been used effectively in the past few years in several areas of medical diagnosis. One use has been for identifying heart defects (8). The applicability of decision tables to medical diagnosis stems from the fact that the diagnostic process is primarily logical rather than computational. Decision tables are an ideal means for expressing complex logical relations between symptoms and diagnosis in a compact and readily understandable form.

The decision table approach is a straightforward approach. Each decision table is divided into four quadrants as in Figure 5. The upper quadrant contains a series of conditions or questions which are to be tested. The lower left quadrant describes the action to be taken depending on the outcome of the tests.

Miscellaneous Approaches

Several other statistical techniques have also been used in diagnosing. One of these techniques is discriminant analysis. Overall (9) has applied the technique to the

study of psychiatric diagnosis as a means of increasing the objectivity and the reliability of classification procedures for psychiatric patients.

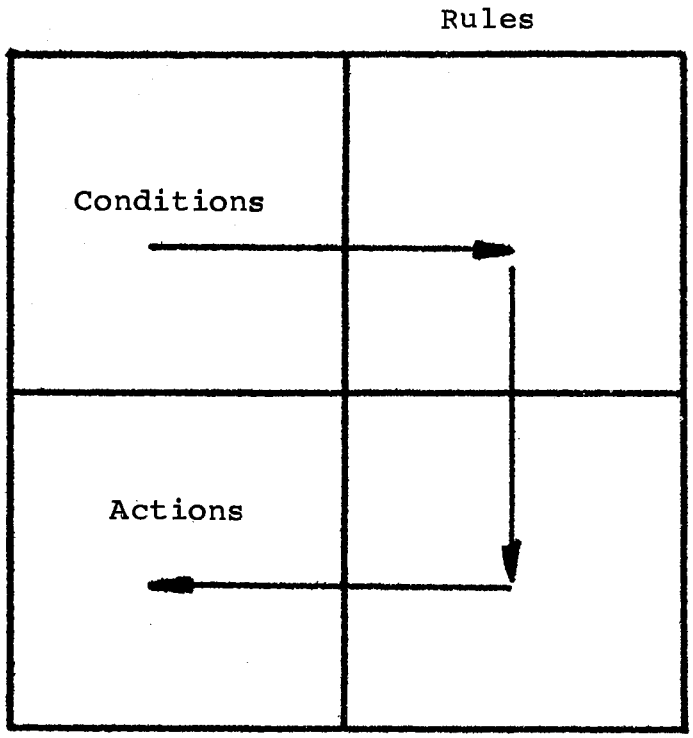


Figure 5. Typical Decision Table Used in Diagnosing

Another statistical technique used in medical diagnosis is factor analysis. Again Overall and Williams (10) have applied the technique to the study of thyroid function diagnosis. One hundred and sixty cases were selected at random for the study and twenty-seven measures of thyroid activity were selected for consideration in the analysis.

Uniqueness of Proposed Approach

All of the approaches discussed in the preceding sections have one thing in common: each approach is oriented toward diagnosing specific diseases or abnormalities. Therefore, each approach relies heavily on obtaining a valid sample of patients, who possess the required diagnoses and/or symptoms, to develop an approach capable of detecting the disease or abnormality.

On the other hand, the approach proposed in this research is to develop a generalized technique which only classifies a patient as normal or abnormal. Such an approach has value as a screening tool for rapidly and economically determining, in general terms, if a patient is healthy or has some abnormality. Should the patient be classified as abnormal (based on a decided risk factor), the patient could be subjected to more thorough testing and/or referred to a specialist for treatment.

CHAPTER II

APPROACH TO THE PROBLEM

This chapter presents the approach to the problem. Included in this chapter are a statement of the problem, the constraints placed on the problem, the source of data, the selected clinical variables, and a discussion of the normality existing in physical variables.

Statement of the Problem

The medical diagnostic process may be divided into the following four general steps:

1. Review of the patient's medical record
2. Comparison of the patient's information with available medical information
3. Diagnosis of the patient
4. Treatment of the patient.

The first step, the review of the patient's medical record, involves the physician familiarizing himself with the patient's medical record. This familiarization consists of reviewing the patient's history data, such as

family disease history, present medications, and smoking habits; his physical examination and lab results; and other procedures, such as chest x-ray and electrocardiogram results.

The second step involves the physician comparing the patient's medical information (i.e., his signs and symptoms) with available medical information. The physician may need to review medical literature or recent medical knowledge to determine the possible causes of signs and symptoms which he has not frequently seen or to reassure himself that he has indeed made the correct diagnosis. One area of assistance to the physician in reviewing available medical literature is the Medical Library Automated Retrieval System (MEDLARS) which is being developed by the National Library of Medicine. MEDLARS is a computer based medical information and retrieval system with remote terminal capabilities.

The third step in the medical diagnosis process is the actual patient diagnosing. It is very difficult, if not impossible, to write a computer program which performs a complete diagnosis. Care must be taken to separate research on the diagnostic processes of the physician from research on the computer techniques which give a medical diagnosis. The diagnostic process for the physician may involve a pattern recognition procedure in the first stage of

diagnosing in order to focus quickly on a group of possible diseases (i.e., the differential diagnosis). The second stage involves the physician using his memory of statistics (i.e., subjective probabilities) to arrive at a final diagnosis.

The first stage of the diagnostic process, the procedure used to arrive at a list of possible diseases, is a difficult area to study. It is very difficult to determine how the physician arrives at a differential diagnosis. Consequently, it is even more difficult to develop a computer program which will perform a differential diagnosis in the sense that, when a new patient is presented, all nonpertinent diseases are screened out and a list of pertinent diseases is retained.

The second stage of the diagnostic process involves the physician making a diagnosis from a given set of diseases. It is this second stage where various mathematical and statistical techniques, using computers, have been applied to arrive at a medical diagnosis. Several of the applications of statistical techniques to medical diagnosis are briefly discussed in Chapter I.

The fourth step in the medical diagnosis process is the actual treatment of the patient.

The research in this thesis is focused on the first three steps in the medical diagnosis process. Emphasis is placed on developing a computerized technique to assist the physician in analyzing a patient's health. No attempt is made to diagnose specific diseases. The primary source of input is the patient's physical examination. Using the physical examination as a basis, the computerized technique will assist the physician by 1) summarizing the patient's prior medical data, 2) determining if the patient is normal or requires additional attention, 3) identifying those clinical variables which are significantly affecting the patient's health, and 4) observing any significant longitudinal drift in the patient's health.

Since emphasis is placed on the physical examination as the primary source of data input, the automated technique for analyzing a patient's health is best suited for occupational health centers, such as industrial medical centers, and for clinics which operate various screening programs.

By emphasizing the physical examination as the primary source of input, the following quantitative variables are available: physical characteristics, vision, hearing, blood specimen, urine specimen, chest x-ray, electrocardiogram, and pulmonary functions.

Imposed Problem Constraints

To adequately limit this research, certain constraints are placed on the problem. These constraints are defined in the following paragraphs.

The first constraint is placed on the selection of the clinical variables. Only data from a patient's physical examination and laboratory tests is considered in this study and then only those variables which are quantifiable. This constraint immediately rules out such variables as psychological variables and other qualitative variables.

A second constraint is imposed because of the availability of data. Because of the researcher's association with the George C. Marshall Space Flight Center's (MSFC) Medical Center, this research is limited to the availability of data from that Center. This constraint should in no way hinder the research. Should the research done on the available data from the Medical Center prove meaningful, it would be just a matter of expanding the number of variables and obtaining an additional data source. The basic techniques should not change.

Because of this second constraint, the majority of the available data is only on males. This is because the majority of employees at MSFC are males. Therefore, only males

are considered in this research. This constraint is not a serious constraint since the model could readily be expanded to include females provided sufficient data is available. In addition, the majority of the data is on employees between twenty-one and sixty-five years old. Therefore, only these ages are considered in this research.

Source of Data

The source of data for this study is the MSFC Medical Center which is located in Huntsville, Alabama. Staffing of the Medical Center consists of five medical doctors and the necessary support personnel. The Medical Center also has its own laboratory.

The primary function of the Medical Center is to oversee the health of MSFC employees. As part of this effort, the Medical Center administers physical examinations and emergency treatment to MSFC employees. In addition, the Center also operates special screening, monitoring, and hazardous occupation programs.

The majority of the data which is generated as a result of the Center's function is entered and stored in the computer. A generalized flow diagram depicting the data input flow is presented in Figure 6.

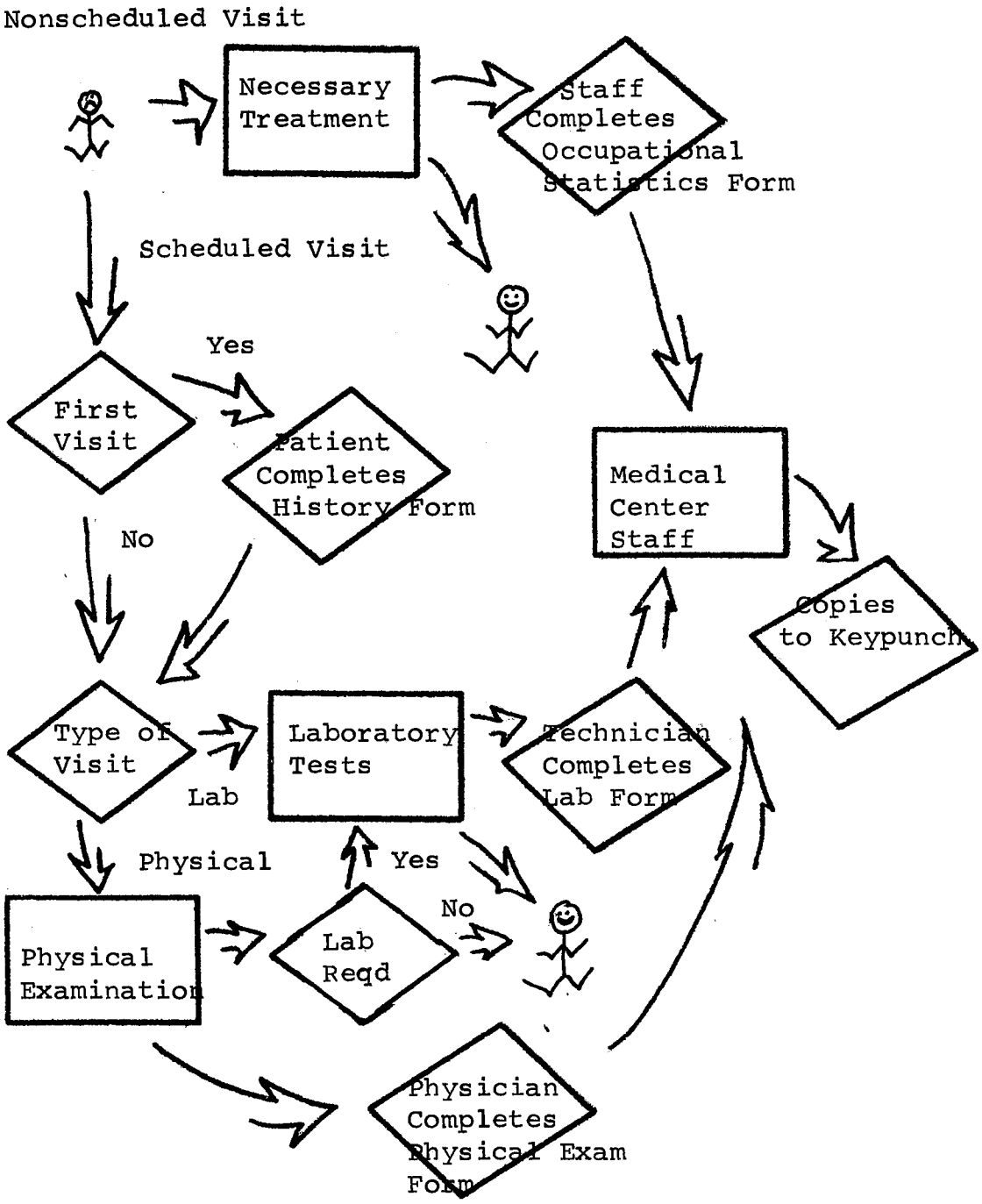


Figure 6. Data Input Flow

When a patient visits the Medical Center one or more of the following forms are completed, depending on the type of visit: history form, job status form, physical examination form, laboratory form, and occupational statistics form.

The history form is completed by the patient only upon his first visit to the Center. The job status form is completed by Medical Center staff whenever there is a change in the employee's job status.

The physical examination form is completed by the examining physician at the time of the patient's physical. The laboratory form is completed by the technician after the laboratory tests have been conducted. The occupational statistics form is completed by the Medical Center staff whenever an employee visits the Center. This includes non-scheduled visits such as emergencies and accidents as well as scheduled visits such as routine physicals.

After the forms have been completed, the Medical Center staff conceal the patient's name on the form, assign a unique medical number to the patient (reassigns the identical number if the patient has previously been assigned a number), and then reproduce the form. By having a unique medical number rather than the patient's name, the patient's medical data is maintained confidential.

All copied forms are forwarded to keypunch. The punched cards are added to the medical data base on a weekly basis. A generalized flow diagram of the updating procedure is presented in Figure 7.

The Medical Center started storing medical data in the computer in 1968. To date, the data base contains data on over 6,000 employees. Of the 30,000 records in the data base, 6,000 are histories, 12,000 are physicals and labs, and 12,000 are occupational statistics records. The primary use of the data base has been as an information storage and retrieval system.

The source of data for this study is the physical examination forms and the laboratory forms. Physicals are given to employees who are less than thirty-one years old every thirty-six months; to employees between thirty-one and forty-five every twenty-four months; and to employees over forty-five every twelve months.

Selected Clinical Variables

Based on the previously discussed constraints, thirty-six clinical variables are selected for this study. These variables can be categorized into physical variables; variables related to blood; variables related to urine;

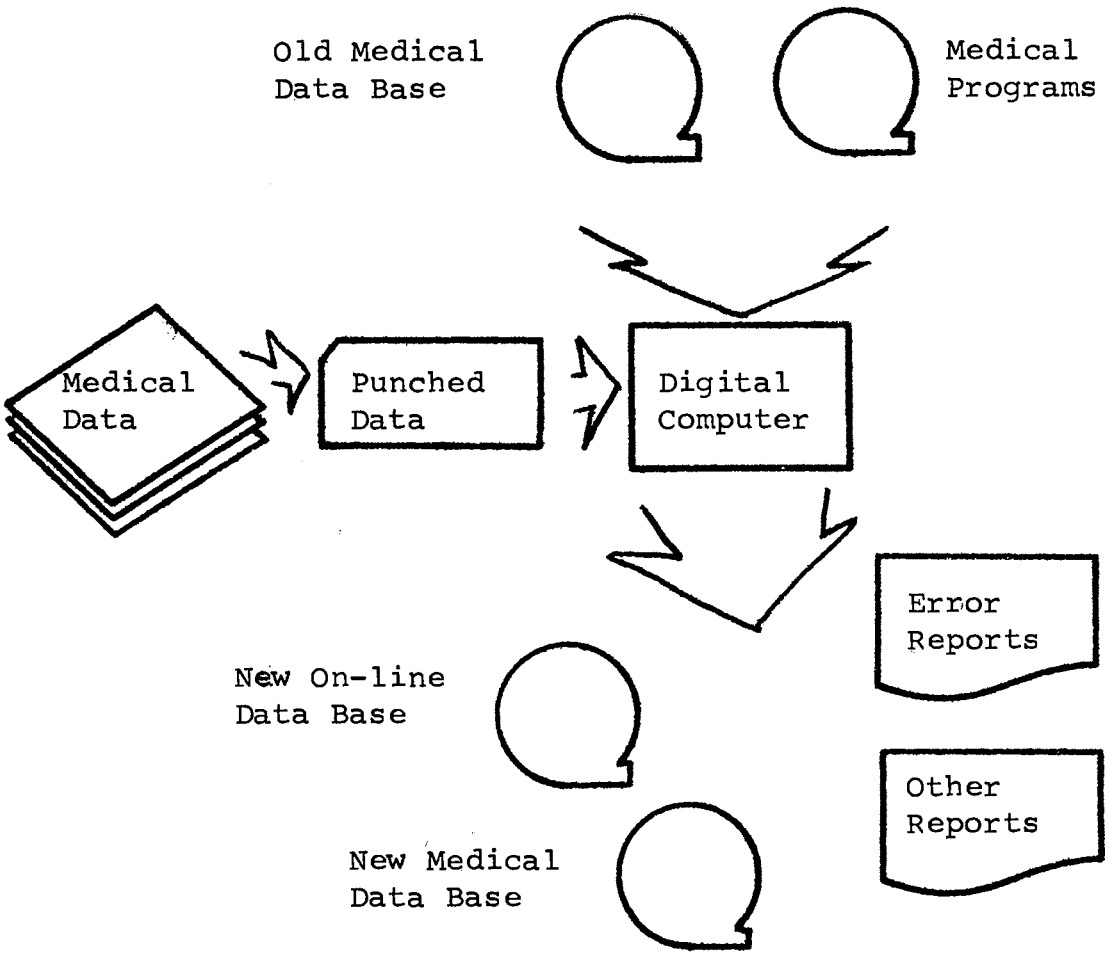


Figure 7. Storing of Medical Data

chest x-ray variables; electrocardiogram variables; and pulmonary function variables.

Within each of these categories the following specific clinical variables are selected:

Physical variables

1. Height (inches)
2. Weight (pounds)
3. Actual/Ideal weight
4. Pulse (beats/minute)
5. Systolic blood pressure (mm Hg)
6. Diastolic blood pressure (mm Hg)
7. Recumbent systolic blood pressure (mm Hg)
8. Recumbent diastolic blood pressure (mm Hg)
9. Arm skin folds (mm)
10. Back skin folds (mm)

Variables related to blood

11. Bilirubin, total (mg/100 ml of serum)
12. Bilirubin, direct (mg/100 ml of serum)
13. Bilirubin, indirect (mg/100 ml of serum)
14. White cell count (number/cu mm)
15. Cholesterol (mg/100 ml of serum)
16. Glucose, fasting (mg/100 ml of serum)
17. Glucose, two hour fasting (mg/100 ml of serum)
18. Hematocrit (% of blood volume)
19. Serum glutamic pyruvic transamenase (SGPT) (units/liter)
20. Thymol turbidity
21. Uric acid (mg/100 ml of serum)

Variables related to urine

22. White blood cells (number/high power field)

23. Red blood cells (number/high power field)
24. Albumin
25. pH
26. Specific gravity
27. Sugar

Chest x-ray variables

28. Total heart diameter (cm)
29. Thoracic diameter (cm)
30. Total heart diameter/thoracic diameter

Electrocardiogram variables

31. Heart rate (beats/minute)
32. PR interval (hundreths of second)
33. QRS duration (hundreths of second)
34. QRS axis (degrees)

Pulmonary function variables

35. Vital capacity (liters)
36. Forced expiration volume in one second (FEV_1) (liters).

A brief description of these clinical variables is presented in Appendix A.

A review of the data availability for the selected variables indicated that recumbent systolic and diastolic blood pressures are only taken when the patient has an elevated blood pressure. Therefore, these two variables are removed from the list. In addition, only total bilirubin is being recorded without recording the direct and indirect bilirubins. Therefore these two variables are removed from the list.

Glucose two hour fasting and thymol turbidity are presently not being measured; therefore, these two variables are also removed from the list. Albumin and sugar from the urine analysis are also not being measured; therefore, these two variables are also removed.

After removing the above variables, twenty-eight variables remain for use in developing the model. It should be noted that once the model is developed, the adding of additional variables would be no major problem.

Phenomenon of Normality

A phenomenon present with clinical variables as with many physical variables is that the distributions of these variables are generally normal. Therefore, these distributions are completely described by knowing their means and standard deviations.

When considering clinical variables we are generally interested in many variables which are not necessarily independent; but which are correlated with one another. Therefore, the joint distribution of these clinical variables is a multivariate normal distribution (11) written as

$$f(\underline{x}) = \frac{1}{2^{n/2} |\underline{V}|^{1/2}} \exp\left[-1/2\ (\underline{x} - \underline{\mu})' \underline{V}^{-1} (\underline{x} - \underline{\mu})\right], \quad (2)$$

where $\underline{x}$ = is a vector of measurements $(x_1, x_2, \ldots, x_p)$,

$\underline{\mu}$ = is a vector of means $(\mu_1, \mu_2, \ldots, \mu_p)$, and

$\underline{V}$ = covariance matrix $\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & & & \vdots \\ \sigma_{31} & & \ddots & & \\ \sigma_{p1} & \cdots & & & \sigma_p^2 \end{bmatrix}$.

The quadratic form of the exponent of the multivariate normal density specifies the equation of an ellipsoid in the p-dimensional space when it is set equal to some positive constant c. Mathematically, this ellipsoid is expressed as

$$(\underline{x} - \underline{\mu})' \underline{V}^{-1} (\underline{x} - \underline{\mu}) = c . \qquad (3)$$

This ellipsoid defines a tolerance region in the p-dimensional space. In effect, by varying c, the "size" of the ellipse is varied. Therefore, c can be considered analogous to defining a tolerance region.

If the values $(x_1, x_2, \ldots, x_p)$ of a patient's clinical variables are known, it is possible to construct a test, with a confidence based on a function of the constant c, to determine if the p-variable observation in the p-dimensional space falls within the ellipsoid. The actual development of this test is in a later chapter.

If the rank of the multivariate normal density is two, the density reduces to a bivariate normal which can be

written as

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-1/2\left[\frac{1}{1-\rho^2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\}. \quad (4)$$

A plot of the bivariate density is shown in Figure 8. If $z_1 = (x_1 - \mu_1)/\sigma_1$ and $z_2 = (x_2 - \mu_2)/\sigma_2$, then $f(x_1,x_2)$ becomes a standardized bivariate density with means zero and unit variances. Mathematically, this is

$$f(z_1,z_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-1/2\,\frac{1}{1-\rho^2}\,(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right] . \quad (5)$$

Any vertical plane in Figure 8 cuts the curve into a univariate normal distribution. Those vertical plane cuts shown in Figure 8 are conditional normal distributions $f(x_1|x_2)$ and $f(x_2|x_1)$.

The curves formed by the intersection of the distribution surface and a horizontal plane is an ellipse as shown in Figure 4. The shape of this ellipse is a function of the correlation coefficient ρ and the relative values of σ_1 and σ_2. The ellipse has its center at μ_1 and μ_2. If $\rho = 0$ and if $\sigma_1 = \sigma_2$, the ellipse becomes a circle.

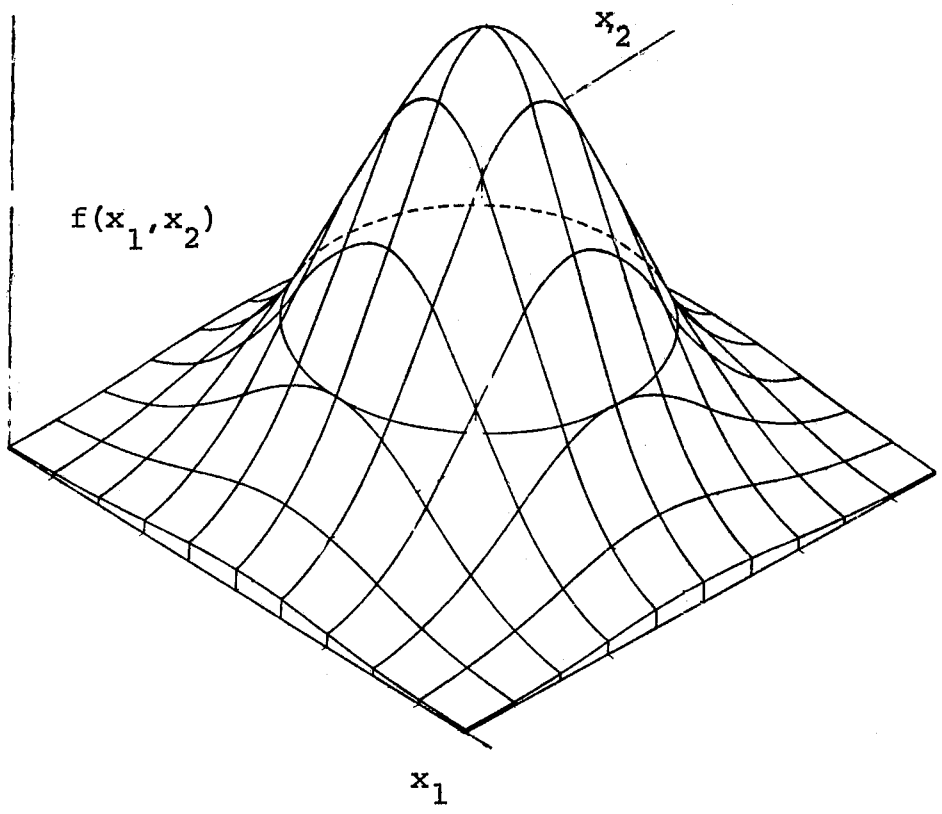


Figure 8. Plot of the Bivariate Normal Density

CHAPTER III

DEVELOPMENT OF THE MODEL

Chapter III presents the development of the model for analyzing a patient's health. Included in this chapter are the classification of the data, the check of the univariate densities for normality, the check of the multivariate densities for normality, the patient classification procedure, the major contributors to the chi-square, the model outputs, and the computer programming and hardware requirements.

Classification of Data

To begin the development of the model, it is assumed that the norms (i.e., norms being the mean vector and covariance matrix) describing a person will vary with time (i.e., a person's age). Therefore, to account for any possible variation, the population is divided into age groups. In effect, age is considered an independent variable: it being the only variable which affects the other variables, but which is not affected by the others.

As previously stated in the imposed constraints, only a small portion of the population is less than twenty-one and older than sixty-five. Therefore, the following age groups are selected: 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, and 60-64.

In collecting the data only employees who had physicals between January, 1970, and October, 1971, are considered. For those employees who, because of their age, had more than one physical during this time period, only their last physical is considered.

The distribution of employees within the nine age groups is given in Table I. Because of the relative small number of patients in the first two age groups, these groups are combined into one age group, thus giving eight age groups. The sample sizes of the variables within each of the eight age groups are given in Table II.

Check of Univariate Densities for Normality

Given a set of p-dimensional observations $(x_{11}, x_{12}, \ldots, x_{1p})$, $(x_{21}, x_{22}, \ldots, x_{2p})$, $\ldots$, $(x_{n1}, x_{n2}, \ldots, x_{np})$, the problem is to determine if the joint probability density function $f(x_1, x_2, \ldots, x_p)$ is normally distributed. To check the joint density for normality, the marginal distributions $f(x_i)$ must be normally distributed. If the

TABLE I

DISTRIBUTION OF PEOPLE WITHIN AGE GROUPS

Age Group	Number of People
20-24	15
25-29	81
30-34	444
35-39	655
40-44	654
45-49	819
50-54	644
55-59	318
60-64	195
Total	3825

TABLE II

DISTRIBUTION OF VARIABLES WITHIN AGE GROUPS

Variable	Age Group								
	20-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	Total
Height	92	216	644	639	805	634	311	155	3496
Weight	92	216	645	639	805	634	312	155	3498
Actual/Ideal Weight	91	212	634	630	795	629	308	154	3453
Pulse	91	214	637	626	791	626	300	149	3434
Systolic Blood Press	92	215	644	638	801	635	309	155	3489
Diastolic Blood Press	92	215	644	638	801	635	309	155	3489
Arm Skin Folds	50	86	244	220	436	349	171	77	1633
Back Skin Folds	40	87	245	220	429	347	170	77	1615
Hematocrit	95	224	654	650	819	644	315	155	3556
White Blood Count	95	224	654	649	819	644	315	155	3555
Glucose	95	216	647	642	811	639	314	154	3518
Cholesterol	95	218	647	641	812	640	314	155	3522
Uric Acid	95	216	645	642	812	639	314	155	3518
SGPT	87	199	600	599	739	562	279	138	3203
Bilirubin, Total	96	224	655	654	819	644	318	156	3566
Urine Red Cells	55	106	314	299	530	431	205	84	2024
Urine White Cells	55	106	314	299	530	431	205	84	2024
Specific Gravity	96	221	648	643	806	632	307	155	3508
Urine PH	96	222	651	649	814	636	310	155	3533
Total Heart Diameter	49	88	256	236	448	354	177	78	1686
Thoracic Diameter	49	88	256	236	448	354	177	78	1686
T.D./Th.D.	49	88	256	236	448	354	177	78	1686
EKG Heart Rate	93	214	636	633	802	636	311	155	3480
PR Interval	49	87	256	231	447	359	178	77	1684
QRS Duration	49	87	255	231	447	358	178	77	1684
QRS Axis	89	194	572	565	666	507	241	105	2939
Vital Capacity	23	24	87	98	242	159	88	45	766
Forced Expiration	23	24	87	98	242	159	88	45	766
Sample size	96	444	655	654	819	644	318	195	3825

marginal distributions are not normally distributed, then $f(\underline{x})$ cannot be normal. However, if $f(x_i)$ is not normal, it may be possible to find a transformation of the variable such that the transformed values are normally distributed.

Graphical Approach

A graphical approach is initially used to assist in checking the marginal distributions for normality. The observations for each clinical variable for each age group are first ranked in ascending order. The expected cumulative probability is computed as

$$\frac{\text{rank number}}{\text{sample size} + 1} . \qquad (6)$$

Each value is then plotted against its expected cumulative probability on normal probability paper. From these plots the appropriate transformations are chosen and the data replotted. In the majority of instances the data is replotted on log normal probability paper. Sample plots of the distributions are given in Appendix B. From the plots of the cumulative distributions, the transformations given in Table III are necessary to make the appropriate $f(x_i)$ normal.

TABLE III

TRANSFORMATION OF VARIABLES

x_i	Variable	Transformation
x_1	Height	None
x_2	Weight	None
x_3	Actual/Ideal Weight	$\text{Log}_e(x_3)$
x_4	Pulse	$\text{Log}_e(x_4)$
x_5	Systolic Blood Pressure	$\text{Log}_e(x_5)$
x_6	Diastolic Blood Pressure	$\text{Log}_e(x_6)$
x_7	Arm Skin Folds	$\text{Log}_e(x_7)$
x_8	Back Skin Folds	$\text{Log}_e(x_8)$
x_9	Hematocrit	None
x_{10}	White Blood Count	$\text{Log}_e(x_{10})$
x_{11}	Glucose	$\text{Log}_e(x_{11})$
x_{12}	Cholesterol	$\text{Log}_e(x_{12})$
x_{13}	Uric Acid	$\text{Log}_e(x_{13})$
x_{14}	SGPT	$\text{Log}_e(x_{14})$
x_{15}	Bilirubin, Total	$\sqrt{x_{15}} + \sqrt{x_{15} + 1}$
x_{16}	Urine Red Cells	$\sqrt{x_{16}} + \sqrt{x_{16} + 1}$
x_{17}	Urine White Cells	$\sqrt{x_{17}} + \sqrt{x_{17} + 1}$
x_{18}	Urine Specific Gravity	None
x_{19}	Urine pH	$\sqrt{x_{19}}$
x_{20}	Total Heart Diameter	None
x_{21}	Thoracic Diameter	None
x_{22}	T.D./ Th.D.	None
x_{23}	EKG Heart Rate	$\text{Log}_e(x_{23})$
x_{24}	PR Interval	$\text{Log}_e(x_{24})$
x_{25}	QRS Duration	$\text{Log}_e(x_{25} + 10)$
x_{26}	QRS Axis	None
x_{27}	Vital Capacity	None
x_{28}	Forced Expiration Volume	None

After plotting the data on probability paper, the data is edited to remove outliers. The rejection of an extreme value (i.e., an outlier) has several possible consequences. One is that the extreme value is due to a faulty observation. The second is that some specific cause has given rise to the extreme value.

The Chauvenet (12) criterion is used for accepting or rejecting extreme values. For a given initial distribution two theoretical values $\tilde{x}_1$ and $\tilde{x}_n$ are defined by

$$n\, F(\tilde{x}_1) = 1/2, \text{ and} \tag{7}$$

$$n\left[1 - F(\tilde{x}_n)\right] = 1/2. \tag{8}$$

From these equations, given n, $F(\tilde{x}_1)$ and $F(\tilde{x}_n)$ are computed. These values establish a set of "limits" which are used in rejecting outliers. Since the distributions are normally distributed, or have been transformed to approximate the normal, these values correspond to the z-values in the normal distribution with mean zero and variance one. Therefore, by knowing z, the limits for accepting values are defined by

$$\mu - z\left[F(\tilde{x}_1)\right]\sigma \text{ , and} \tag{9}$$

$$\mu + z\left[F(\tilde{x}_n)\right]\sigma \text{ .} \tag{10}$$

As a result of using the Chauvenet criterion, 501 outliers are deleted. A distribution of these outliers by age group is given in Table IV.

TABLE IV

DISTRIBUTION OF UNIVARIATE OUTLIERS

Age Group	Number of Outliers	Percent of Sample
20-29	30	1.8
30-34	41	1.2
35-39	72	0.7
40-44	67	0.7
45-49	106	0.7
50-54	84	0.7
55-59	71	1.2
60-64	30	1.1
Total	501	0.8

Statistical Check for Normality

In addition to plotting each variable on the appropriate probability paper, and after the outliers are removed, a more rigorous test for normality is performed. The test is the Kolmogorov-Smirnov (K-S) test (13). This test is based on the fact that the observed cumulative distribution

of a sample is expected to be fairly close to the theoretical distribution. The goodness of fit is measured by finding the point at which the observed distribution and the theoretical distribution are the farthest apart and then comparing this distance with the entry in a table of critical values which indicate whether such a large distance is likely to occur.

If the distance is too large, the chance that the distribution is actually the theoretical distribution is very small. Mathematically, the K-S test is

$$\text{Max} \left[F(x)_{\text{theo}} - F(x)_{\text{actual}} \right] < \text{K-S value.} \qquad (11)$$

For sample sizes greater than 50 and a level of significance of 0.05, the K-S value is $1.36/\sqrt{n}$. For a level of significance of 0.10 the K-S value is $1.63/\sqrt{n}$.

The results of the K-S tests are given in Appendix C. The means in Appendix C for those variables which are transformed, are the transformed means. For example, the actual systolic blood pressure for age group 20-29 is the natural log of 4.71612 which is 112. Tables V and VI contain the actual means and standard deviations for the age groups.

TABLE V

AGE GROUP MEANS

Clinical variable	Age group 20-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64
HEIGHT	70.780	71.209	71.066	70.635	70.372	70.415	70.188	69.693
WEIGHT	169.250	176.952	180.167	179.038	179.025	178.755	179.716	178.040
ACTUAL WEIGHT/IDEAL WEIGHT	1.082	1.127	1.137	1.143	1.153	1.151	1.167	1.169
PULSE	71.977	73.478	73.506	74.693	74.643	74.679	74.804	74.842
SYSTOLIC BLOOD PRESSURE	112.483	112.047	113.473	115.874	118.297	121.115	126.690	129.075
DIASTOLIC BLOOD PRESSURE	69.282	70.516	72.603	73.830	75.816	76.627	78.233	77.678
ARM SKIN FOLDS	10.553	11.963	11.594	12.157	11.913	11.919	11.717	10.507
BACK SKIN FOLDS	15.133	16.734	17.586	18.447	18.358	17.756	18.276	17.647
HEMATOCRIT	45.042	45.032	44.845	45.026	45.128	45.296	45.169	44.904
WHITE BLOOD COUNT	65.659	67.790	68.464	72.916	71.078	72.762	71.950	71.585
GLUCOSE FASTING	82.319	84.892	84.476	86.639	84.954	87.356	87.964	90.710
CHOLESTEROL	208.736	220.446	227.855	235.759	239.731	241.317	240.732	240.168
URIC ACID	5.043	5.023	4.996	5.072	4.981	4.962	5.039	5.038
SGPT	14.069	15.082	14.979	14.767	13.361	13.086	12.292	11.882
TOTAL BILIRUBIN	3.021	2.612	3.046	3.146	3.181	3.185	3.176	3.744
URINE RED CELL COUNT	.036	.019	.036	.000	.023	.028	.010	.103
URINE WHITE CELL COUNT	1.370	1.990	1.409	1.244	1.316	1.325	1.426	1.649
URINE SPECIFIC GRAVITY	1.017	1.016	1.016	1.016	1.015	1.016	1.015	1.015
URINE PH	6.000	5.904	5.925	5.925	5.897	5.850	5.849	5.976
TOTAL HEART DIAMETER	12.428	12.505	12.959	13.360	13.366	13.505	13.671	13.552
THORACIC DIAMETER	30.521	31.394	31.420	31.548	31.248	31.444	31.532	31.142
TOTAL HEART DIA/THORACIC DIAMETER	.407	.398	.412	.424	.428	.430	.433	.435
EKG HEART RATE	63.565	65.933	64.908	66.604	66.790	67.312	67.462	65.646
PR INTERVAL	15.729	15.630	15.983	16.009	16.379	16.488	16.520	16.529
QRS INTERVAL	7.673	7.590	7.586	7.495	7.500	7.582	7.747	7.782
QRS AXIS	57.867	50.520	44.406	40.000	34.628	30.386	28.487	20.946
VITAL CAPACITY (VC)	4.940	5.004	4.773	4.608	4.415	4.362	4.043	3.805
FORCED EXPIRATION VOLUME (FEV1)	4.145	4.066	3.849	3.634	3.432	3.354	3.104	2.908

TABLE VI

AGE GROUP STANDARD DEVIATIONS

Clinical variable	Age group 20-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64
HEIGHT	2.502	2.330	2.377	2.447	2.325	2.355	2.355	2.319
WEIGHT	24.160	22.611	23.886	23.554	24.083	23.901	23.516	23.079
ACTUAL WEIGHT/IDEAL WEIGHT	.133	.127	.130	.129	.137	.136	.135	.132
PULSE	8.215	9.128	9.455	9.632	9.349	9.257	9.757	8.736
SYSTOLIC BLOOD PRESSURE	13.165	12.821	12.569	14.211	16.465	17.420	19.016	18.173
DIASTOLIC BLOOD PRESSURE	9.218	9.168	9.204	9.758	10.819	10.744	10.384	10.210
ARM SKIN FOLDS	5.063	5.301	5.285	5.482	5.501	5.386	5.033	4.457
BACK SKIN FOLDS	6.341	6.289	6.813	6.182	6.682	6.589	5.786	6.410
HEMATOCRIT	2.047	2.316	2.268	2.396	2.506	2.520	2.475	2.196
WHITE BLOOD COUNT	11.842	13.855	14.185	15.487	14.946	16.481	14.359	13.816
GLUCOSE FASTING	9.706	10.682	11.908	13.842	12.394	12.972	12.791	16.752
CHOLESTEROL	38.872	41.215	37.252	40.386	41.089	39.776	39.881	37.298
URIC ACID	.943	1.052	1.121	1.127	1.066	1.104	1.114	1.020
SGPT	8.989	8.843	8.672	9.024	7.153	7.128	5.489	5.417
TOTAL BILIRUBIN	2.645	2.765	2.530	2.704	2.641	2.678	2.522	2.517
URINE RED CELL COUNT	.269	.138	.273	.000	.205	.166	.100	.416
URINE WHITE CELL COUNT	1.137	3.892	1.680	1.277	1.566	1.459	1.428	1.536
URINE SPECIFIC GRAVITY	.005	.006	.005	.006	.005	.005	.005	.006
URINE PH	.695	.753	.756	.777	.821	.793	.832	.800
TOTAL HEART DIAMETER	1.471	1.087	1.379	1.409	1.310	1.335	1.294	1.463
THORACIC DIAMETER	1.517	1.457	1.761	1.822	1.733	1.766	1.769	1.991
TOTAL HEART DIA/THORACIC DIAMETER	.043	.029	.037	.038	.036	.039	.036	.036
EKG HEART RATE	8.999	11.137	10.656	11.154	10.778	11.023	11.262	9.362
PR INTERVAL	1.898	1.893	1.989	1.835	2.108	2.178	2.064	2.255
QRS INTERVAL	.898	.781	.844	.886	.921	.993	1.319	1.293
QRS AXIS	26.698	29.314	29.997	30.436	31.335	33.059	32.550	31.855
VITAL CAPACITY (VC)	.337	.725	.699	.623	.659	.667	.609	.668
FORCED EXPIRATION VOLUME (FEV1)	.386	.604	.555	.509	.563	.587	.579	.621

Check of Multivariate Densities for Normality

If the marginal distributions $f(x_i)$ are normal, or can, through the appropriate transformation, be normal, then the joint distribution $f(\underset{\sim}{x})$ may or may not be normal. Therefore, the joint distribution must be checked for normality.

Each of the age groups in the previous section has its own mean vector and covariance matrix. Also each variable within the age groups is normally distributed. To check each of the joint distributions for normality, probability plots (14) similar to the normal probability plots in the univariate checks are used.

The gamma probability density function is

$$f(y;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{\alpha-1} e^{-y/\beta}, \qquad (12)$$

where α is a shape parameter, and β is a scale parameter.

The gamma distribution function is

$$F(y;\alpha,\beta) = \int_0^y f(y;\alpha,\beta)\,dy. \qquad (13)$$

Let $y_1 \leq y_2 \leq \ldots \leq y_n$ be an ordered random sample of n observations. Let $b_1, b_2, \ldots, b_n$ be appropriate chosen fractions of the gamma distribution corresponding to the y's. If $\tilde{y}_i$, $i = 1,2, \ldots,n$ satisfies

$$F(\tilde{y}_i;\alpha,\beta) = b_i \quad i = 1,2, \ldots, n, \qquad (14)$$

and if the y's are a random sample from a gamma distribution with parameters α and β, then the points $(\tilde{y}_i, y_i)$ $i = 1, 2, \ldots, n$ will fall along a straight line with slope one through the origin.

Using the linear transformation $x = \beta(y - \alpha)$ reduces the gamma distribution to a standard form with $\beta = 1$ to

$$F(x; \alpha, 1) = \int_0^x \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} dx. \qquad (15)$$

If $\tilde{x}_i$ satisfies

$$F(\tilde{x}_i; \alpha, 1) = b_i, \qquad (16)$$

then the plot of the points (x_i, x_i) $i = 1, 2, \ldots, n$ falls along a straight line with slop equal to $1/\beta$.

If $\alpha = p/2$ and $\beta = 1/2$, $f(x; \alpha, \beta)$ is a chi-square distribution with p degrees of freedom. The plot of $(\tilde{x}_i, x_i)$ then falls along a straight line with slope two.

The quadratic form Q(x) of the multivariate normal density function follows a chi-square distribution with p degrees of freedom. Therefore, each patient in the age group sample has a Q(x) value. These Q(x)'s correspond to the $\tilde{x}_i$ in equation (16) and can be plotted using the gamma probability plots for each of the age groups. The gamma plots are given in Appendix D.

The Chauvenet criterion is again used to eliminate outliers from the multivariate normal distributions. As a result of using the Chauvenet criterion, 63 outliers are deleted. The distribution of these outliers for each age group is given in Table VII.

TABLE VII

DISTRIBUTION OF MULTIVARIATE OUTLIERS

Age Group	Number of Outliers	Percent of Sample
20-29	3	0.6
30-34	2	0.4
35-39	10	5.3
40-44	8	4.7
45-49	15	4.2
50-54	11	4.2
55-59	6	4.3
60-64	8	14.1
Total	63	4.9

After removing the multivariate outliers, the Q(x)'s are replotted using the gamma probability plots. These plots are also plotted in Appendix D by the symbol ▲ .

Grouping of Age Groups

It was initially assumed that the norms vary with age; therefore, norms were developed for eight age groups. However, since these age groups are arbitrarily assigned, it may be possible to pool some of the groups. There are several advantages to pooling the age groups. One obvious advantage is a reduction in required computer storage. Another advantage is that, by pooling similar age groups, a larger sample is available to estimate the norms.

The pooling of age groups is divided into two areas. The first area is the pooling of the mean vectors. The second area is the pooling of the covariance matrices.

Pooling of Mean Vectors

Two approaches are used in attempting to pool similar mean vectors. The first approach is cluster analysis while the second approach is a chi-square test.

Cluster Analysis. Cluster analysis is a technique for investigating the relationships of points in a multidimensional space for the purpose of identifying those points which tend to cluster together. Since the age groups are arbitrarily selected, by using cluster analysis, it is possible to identify those age groups which tend to have

similar mean vectors. Therefore, those age groups for which the mean vectors tend to cluster together can be combined into a single age group.

The clustering technique which is used was developed by Edwards and Cavalli-Sforza (15). From the analysis of variance it is known that the sum of the squared distances of points on a line from their mean can be partitioned, when the points are classified into two groups, into two within-groups sum of squares and a between-groups sum of squares. Since all the quantities involved are squared distances, it is evident that this is also true for points in any number of dimensions, because these squared distances can all be partitioned into squared distances along the Cartesian axes, so that if the partition is possible along each axis it is possible among the points as a whole.

Therefore, when points are divided into two clusters, the sum of the squared distances from their mean can be partitioned into the sum of the squared distances of the points of one cluster from their mean, the similar sum for the other cluster, and the between-clusters sum of squares. This is nothing but a single classification analysis of variance conducted in many dimensions. The natural criterion for division is clearly the between-clusters sum of squares, and the best split is that for which this sum is a maximum

and the within-clusters sum of squares consequently a minimum.

Continued splitting according to this criterion will lead to a tree diagram. With each branching will be associated a between-clusters sum of squares, which will be a measure of the importance of the split. Further, since at the end of this process each cluster contains only one point, there is no within-clusters sum of squares left, and the total of the sums of squares associated with each branching must exactly equal the original sum of squares: all the original variation is accounted for.

The Mahalanobis D^2 is used to compute the measure of the distance between paired age group mean vectors. Mathematically,

$$D_{ij}^2 = (\underline{\mu}_i - \underline{\mu}_j)' \underline{V}^{-1} (\underline{\mu}_i - \underline{\mu}_j), \qquad (17)$$

where $\underline{\mu}_i$ and $\underline{\mu}_j$ are the mean vectors for the i^{th} and j^{th} age group, and $\underline{V}^{-1}$ is the inverse of the common covariance matrix of the combined age groups.

The half matrix of the squared distances (i.e., D_{ij}^2) is given in Table VIII. The sum of the D^2 values in Table VIII is 170.233; therefore, the sum of squares is 170.233/8, or 21.278. For example, in investigating the split 1-2-3:4-5-6-7-8, the sum of squares for the 1-2-3

TABLE VIII

D^2 VALUES BETWEEN PAIRED AGE GROUPS

Age Group	Age Group							
	20-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64
20-29		5.636	7.865	10.775	14.058	14.568	21.638	25.222
30-34			0.792	2.233	3.942	4.705	8.821	12.156
35-39				0.546	1.472	2.089	5.237	8.022
40-44					0.543	0.854	3.249	5.807
45-49						0.234	1.564	3.509
50-54							1.136	2.730
55-59								0.820
60-64								

cluster is (5.636 + 7.865 + 0.792)/3, or 4.764. For the 4-5-6-7-8 cluster the sum of squares is 20.446/5, or 4.089. Therefore, the total within-clusters sum of squares is 4.764 + 4.089, or 8.853, and the total between-clusters sum of squares is 21.278 - 8.853, or 12.425. All other possible splits are similarily compared with the best split being the one which maximizes the between-cluster sum of squares.

The best split is found to be 1-2-3:4-5-6-7-8 with a within clusters sum of squares of 8.854 and a between-clusters sum of squares of 12.425. By further splitting these two clusters the final tree diagram in Figure 9 is formed.

The final clustering indicates that several possibilities exist for pooling the mean vectors. For example, age groups 30-34 and 35-39, 45-49 and 50-54, and 55-59 and 60-64 could be pooled. Or, age group 20-29, 30-34, and 35-39 could be pooled.

Table IX gives the within-cluster sum of squares expressed as percentages of the total sum of squares. The sum of the percentages is 26.9 percent; therefore, 73.1 percent of the variation has been accounted for by the arrangement of the tree diagram, which is very successful.

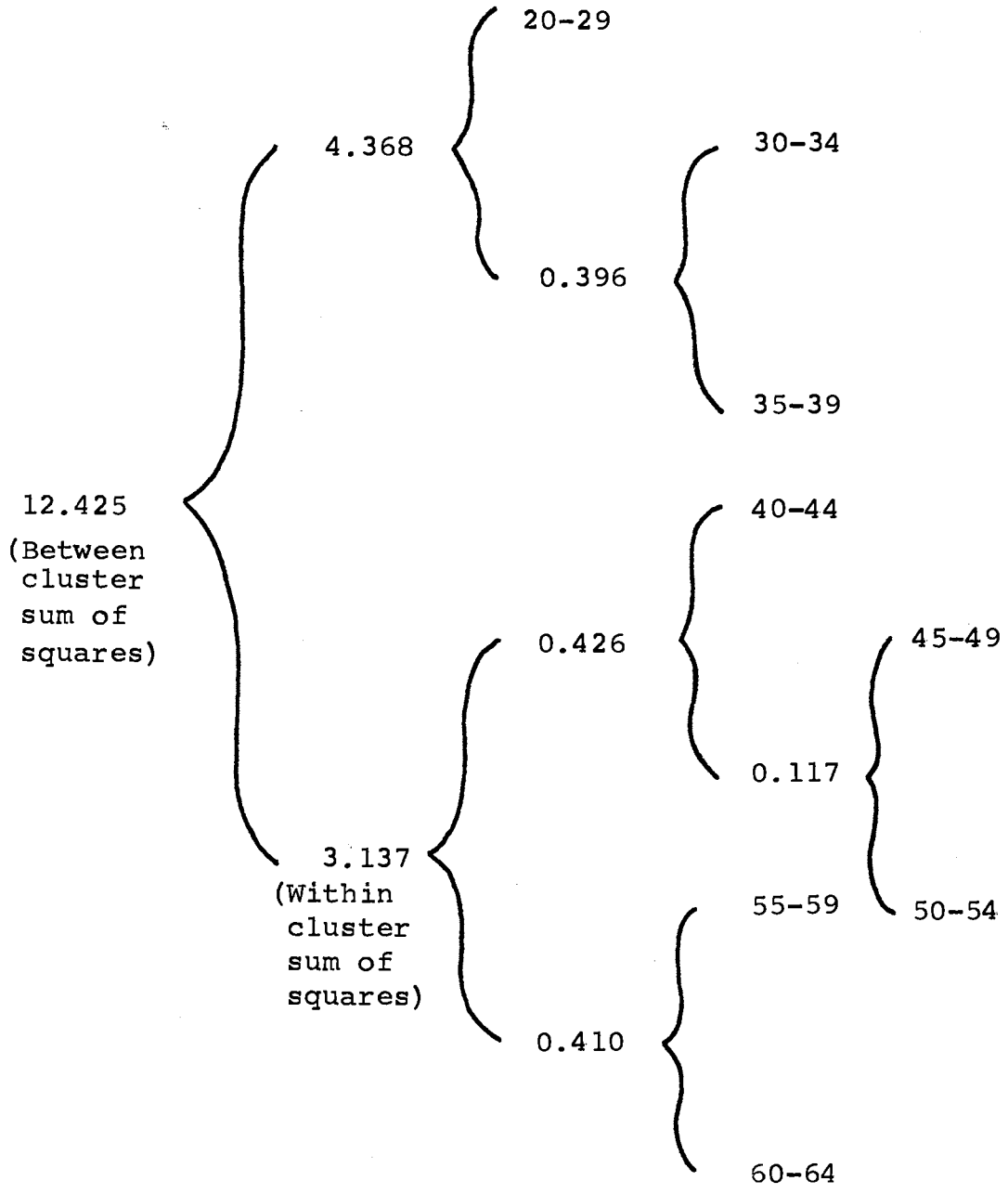


Figure 9. Clustering of Age Groups

TABLE IX

SUMMARY OF CLUSTERING

Age Group	Within-Cluster Sum of Squares	Percent of Total Sum of Squares
20-29, 30-34, 35-39	4.764	22.4
40-44, 45-49, 50-54	0.543	2.6
55-59, 60-64	0.410	1.9
Total		26.9

Chi-Square Test. In testing for a significant difference in the mean vectors, it is assumed that the population covariance is known. Therefore, the chi-square test (16) for the difference of means could be used. Eight age groups have been defined; therefore, twenty-eight possible tests between paired mean vectors are required.

If two samples (i.e., age groups) with mean vectors $\bar{\underline{x}}_i$ and $\bar{\underline{x}}_j$ and sample sizes n_i and n_j respectively, are drawn from a multivariate normal population with a known covariance matrix $\underline{V}$, the following hypothesis can be tested

$$H_o: \underline{\mu}_i = \underline{\mu}_j \tag{18}$$

that the mean vectors are equal by the test statistic

$$\chi^2_{ij} = \frac{n_i\, n_j}{n_i + n_j} \left(\bar{\underline{x}}_i - \bar{\underline{x}}_j\right)' \underline{V}^{-1} \left(\bar{\underline{x}}_i - \bar{\underline{x}}_j\right). \tag{19}$$

If H_o is true, χ^2_{ij} is chi-squared distributed with p degrees of freedom. H_o is accepted at the α level if $\chi^2_{ij} < \chi^2_{\alpha;p}$ and rejected if the statistic exceeds $\chi^2_{\alpha;p}$.

The computed chi-square values between paired age group mean vectors are given in Table X. For $\alpha = 0.05$ and 26 degrees of freedom the theoretical chi-square value is 38.9. Therefore, based solely on the chi-square tests, none of the age group mean vectors can be pooled.

TABLE X

CHI-SQUARE VALUES BETWEEN PAIRED AGE GROUP MEAN VECTORS

Age Group	Age Group 20-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64
20-29		271.72	479.90	655.59	890.85	898.49	1180.45	1103.32
30-34			90.36	253.44	483.85	547.83	823.99	797.95
35-39				122.68	387.38	493.38	824.78	737.42
40-44					141.19	199.59	507.96	531.50
45-49						64.58	272.45	341.65
50-54							183.92	255.03
55-59								63.99
60-64								

$$\chi^2_{(26)\,\alpha=0.05} = 38.9$$

However, the results of the cluster analysis indicate that age groups 45-49 and 50-54 and 55-59 and 60-64 could be pooled. The corresponding chi-square values for these paired age groups are 64.58 and 63.99 which are the lowest values in the table. After discussion with the medical staff it was decided to pool these mean vectors because of the favorable cluster analysis results and the doctor's comments that 45 and 55 are good age breaks. In addition, the pooling of these age groups reduces the required core.

Pooling of Covariance Matrices

A graphical approach (17) is used for testing the age group covariance matrices for homogenity. It is known that

$$(n - 1) \frac{\underline{a}' \hat{\underline{V}} \underline{a}}{\underline{a}' \underline{V} \underline{a}}, \qquad (20)$$

where $\underline{V}$ = population covariance matrix,

$\hat{\underline{V}}$ = an estimate of $\underline{V}$,

n = sample size, and

$\underline{a}$ = an arbitrary vector,

is distributed as the chi-square distribution with n-1 degrees of freedom.

Likewise, values of equation (20) from several independent estimates of $\underline{V}$ would be from the same chi-square

distribution. By observing the distribution of the values, some indication would be given as to the homogenity of the covariance matrix $\underline{V}$. Quite obviously the selection of $\underline{a}$ would influence the sensitivity of the discrimination.

Initially $\underline{a} = (1,1,\ldots,1)$. Then for each of the eight age groups the value $\underline{a}' \hat{\underline{V}}_i \underline{a}$ is computed, where $\hat{\underline{V}}_i$ is an estimate of the covariance for age group i. The value for $\underline{a}' \underline{V} \underline{a}$ is computed as

$$\frac{\sum (n_i - 1)\ \underline{a}' \hat{\underline{V}}_i\ \underline{a}}{\sum\ (n_i - 1)} \quad . \tag{21}$$

In order to graphically display values from equation (20), the chi-square variate is transformed to a normal variate by the transformation

$$u_i = \sqrt{9p/2}\left[(\chi^2)^{1/3} - (1 - 2/9p)\right] \quad , \tag{22}$$

where p = the degrees of freedom (i.e., the rank of $\underline{V}$).

The u_i's are normally distributed with mean zero and variance one. These u_i's are then ranked in ascending order and plotted on normal probability paper. Nonhomogenous age groups should appear as outliers on the plot.

The plots of the u_i's for $\underline{a} = (\underline{1})$ are given in Figure 10. The data for generating Figure 10 is given in Table XI. Since the u_i's are N(0,1), the $(u_i\text{'s})^2$ are chi-squared distributed with eight degrees of freedom. In Table XI the

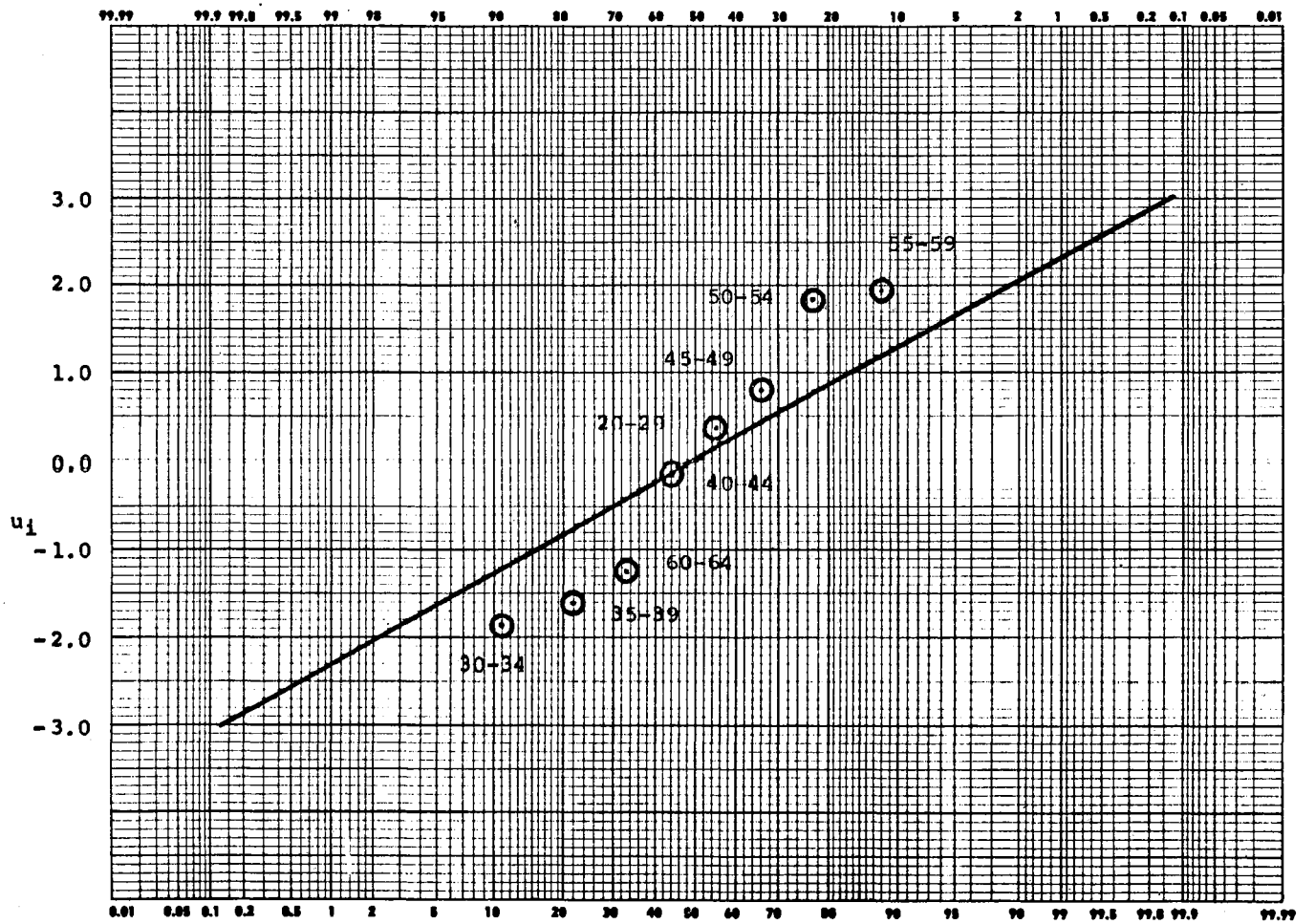


Figure 10. Testing Homogenity of Covariances Using $\underline{a} = (\underline{1})$

sum of u_i^2 = 15.46 which is less than $\chi^2(8)_{\alpha = 0.05}$ = 15.5. Therefore, all the age group covariances could be considered homogenous.

TABLE XI

DATA FOR TESTING HOMOGENITY OF COVARIANCES ($\underline{a}$ = $\underline{1}$)

Age Group	u_i	u_i^2
20-29	0.37	0.14
30-34	-1.87	3.50
35-39	-1.60	2.56
40-44	-0.17	0.03
45-49	0.81	0.66
50-54	1.82	3.32
55-59	1.94	3.76
60-64	-1.22	1.49

However, greater discrimination could be obtained by making $\underline{a}$ equal to the eigenvector of the largest eigenvalue of the pooled covariance matrix. The results are given in Figure 11. The data for generating Figure 11 is given in Table XII. The sum of the u_i^2 = 19.58 which is greater than $\chi^2(8)_{\alpha = 0.05}$ = 15.5. Therefore, the age group covariances are not homogenous.

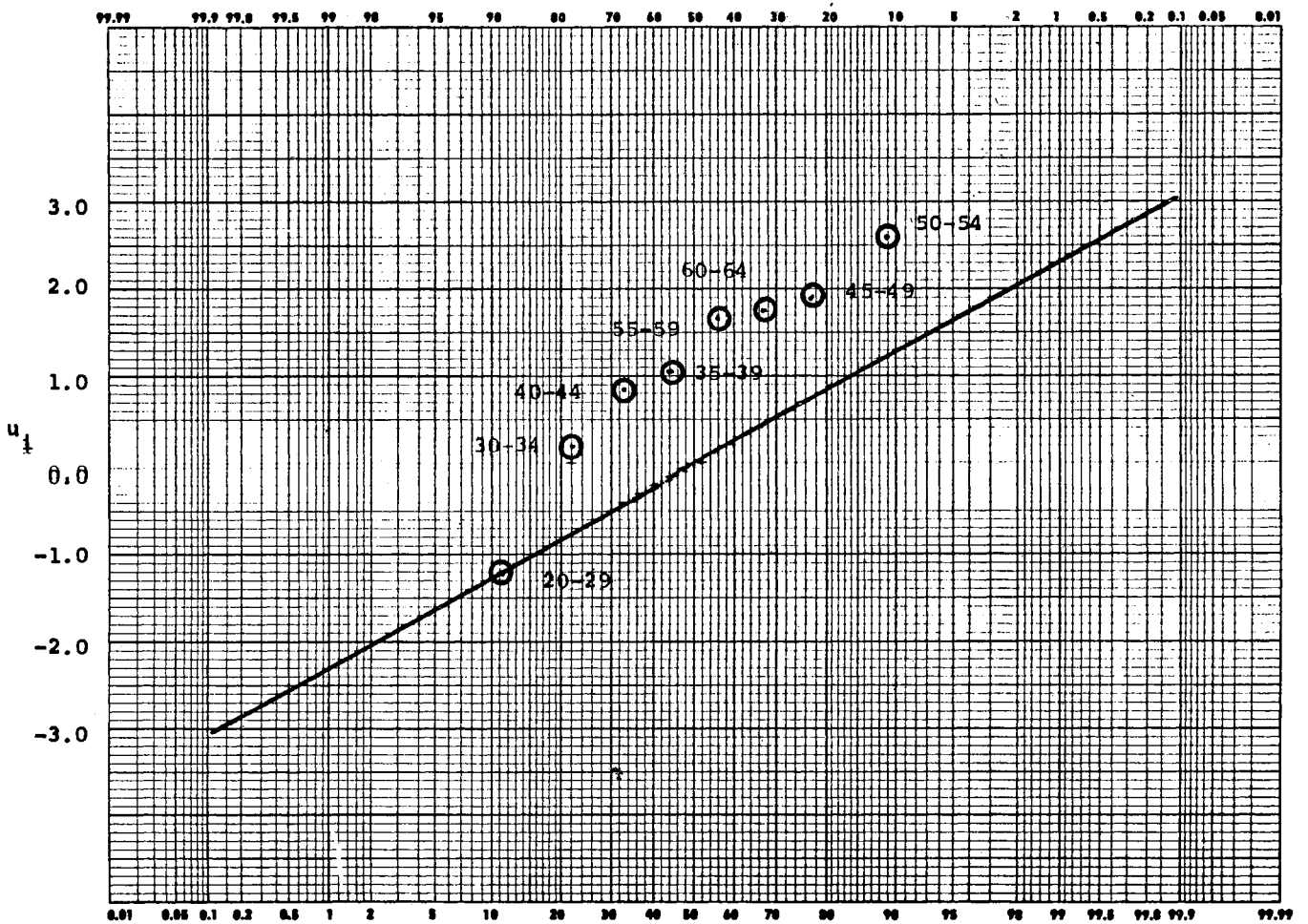


Figure 11. Testing Homogenity of Covariances Using
$\underline{a}$ = Eigenvector of Largest Eigenvalue

TABLE XII

DATA FOR TESTING HOMOGENITY OF COVARIANCES
($\underline{a}$ = EIGENVECTOR OF LARGEST EIGENVALUE)

Age Group	u_i	u_i^2
20-29	-1.21	1.46
30-34	0.21	0.04
35-39	1.08	1.16
40-44	0.84	0.71
45-49	1.91	3.64
50-54	2.60	6.75
55-59	1.63	2.66
60-64	1.78	3.16

Referring to Figure 11 it appears that age groups 20-29 and 30-34 could possibly be pooled as well as several older age groups. Several combinations were run with the best results being the pooling of age groups 50-54, 55-59, and 60-64. The results are given in Figure 12. The data for generating the figure is given in Table XIII.

The sum of u_i^2 for the first five age groups is 7.01 which is less than $\chi^2(5)_{\alpha = 0.05} = 11.1$. The sum of u_i^2 for the last three age groups is 2.55 which is less than $\chi^2(3)_{\alpha = 0.05} = 7.81$; therefore, age 50-64 is homogenous.

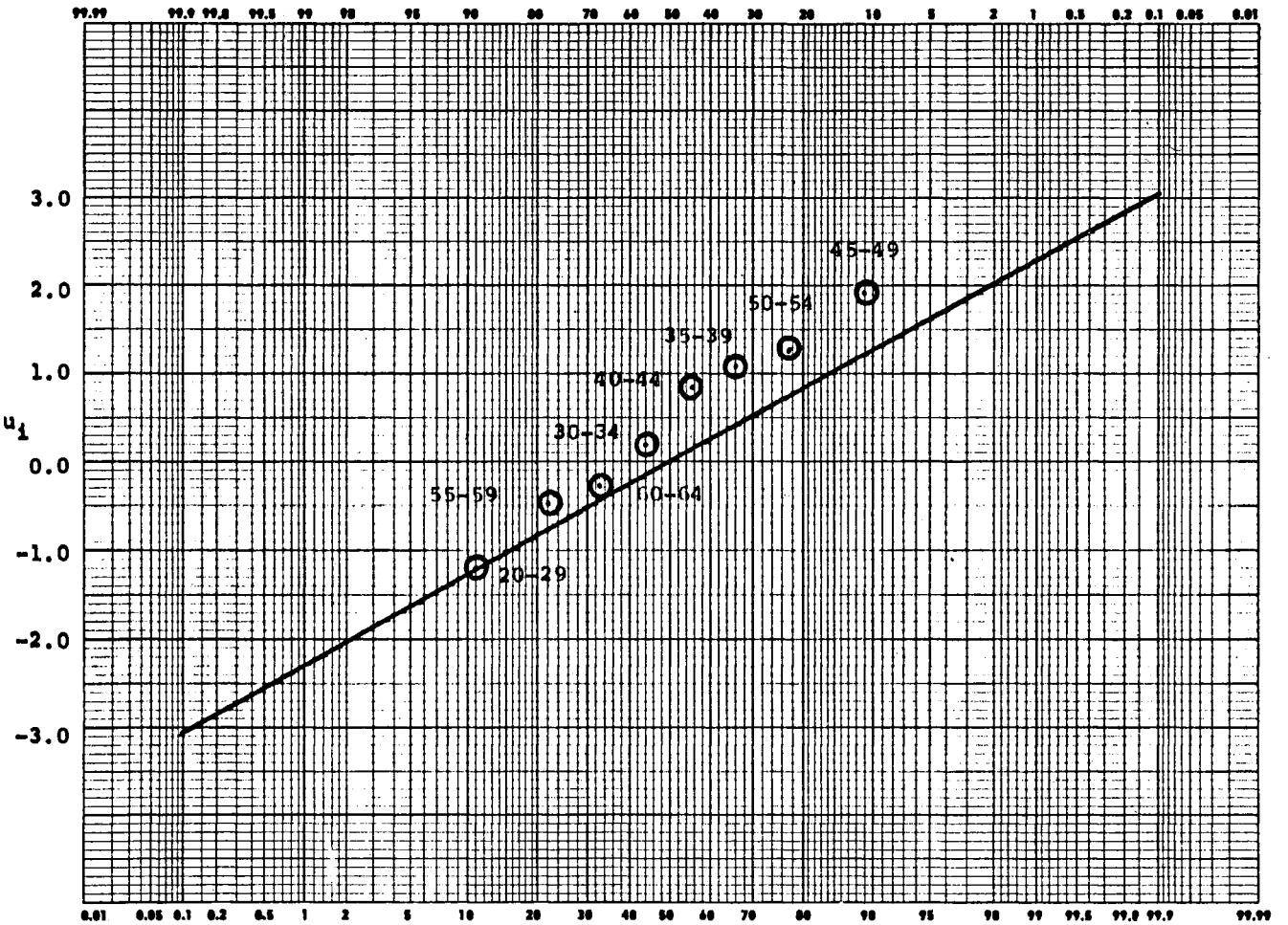


Figure 12. Testing Homogenity of Pooled Covariances Using $\underline{a}$ = Eigenvector of Largest Eigenvalue

TABLE XIII

DATA FOR TESTING HOMOGENITY OF POOLED COVARIANCES
($\underline{a}$ = EIGENVECTOR OF LARGEST EIGENVALUE)

Age Group	u_i	u_i^2
20-29	-1.21	1.46
30-34	0.21	0.04
35-39	1.08	1.16
40-44	1.91	3.64
45-49	1.24	1.54
50-54	-0.48	0.23
60-64	-0.28	0.78

Selected Age Groups

The pooling of the age group mean vectors resulted in the following age group mean vectors: 20-29, 30-34, 35-39, 40-44, 45-54, and 55-64. The new mean vectors for the last two pooled age groups are given in Appendix C.

The pooling of the age group covariance matrices resulted in the following age group covariance matrices: 20-29, 30-34, 35-39, 40-44, 45-49, and 50-64. Listings of these covariance matrices are given in Appendix E.

The multivariate check of age groups 50-54 and 55-64 using the pooled covariance matrix is given in Appendix D.

In general, the pooling of the older age groups did result in a better fit of the data to the multivariate normal.

The correlation between two clinical variables x_i and x_j is defined as

$$\rho_{ij} = \frac{\sigma_{x_i x_j}}{\sigma_{x_i} \sigma_{x_j}}, \tag{23}$$

where $\sigma_{x_i x_j}$ = covariance of x_i and x_j, and

$\sigma_{x_i}\sigma_{x_j}$ = is the standard deviation of x_i and x_j.

The correlations between the clinical variables when all the age groups are combined are given in Table XIV. In testing the hypothesis that $\rho_{ij} = 0$, the test statistic

$$t = \frac{\rho_{ij}\sqrt{n-2}}{\sqrt{1-\rho_{ij}^2}}, \tag{24}$$

is used which has a t-distribution with n-2 degrees of freedom. Those correlations in Table XIV which are greater than $|0.04|$ are significant at the one percent level.

Patient Classification Procedure

The age groups have been defined and the corresponding norms for each age group established. The next step is to develop a procedure for comparing a patient's set of clinical values against the predefined age group norms. Two statistical tests are used for classifying a patient as

TABLE XIV

POOLED CORRELATIONS BETWEEN VARIABLES

	Height	Weight	Actual/Ideal Weight	Pulse	Systolic Blood Press	Diastolic Blood Press	Arm Skin Folds	Back Skin Folds	Hematocrit	White Blood Count	Glucose	Cholesterol	Uric Acid	SGPT	Bilirubin, Total	Urine Red Cells	Urine White Cells	Specific Gravity	pH	Total Heart Dia	Thoracic Dia	T.D./TH.D.	EKG Heart Rate	PR Interval	QRS Duration	QRS Axis	Vital Capacity	Forced Expiration Vol
Height	1.00	.47	-.02	-.04	.04	.04	.04	.05	.00	-.04	-.04	-.06	.01	-.01	-.02	-.00	.02	.01	.01	.12	.37	-.11	-.05	.06	.07	.07	.45	.39
Weight		1.00	.86	.00	.25	.26	.34	.51	.03	-.01	.08	.03	.11	.18	.05	.02	.01	.08	-.02	.54	.56	.26	-.01	.07	.03	-.20	.18	.20
Actual/Ideal Weight			1.00	.02	.26	.27	.38	.58	.03	.01	.12	.07	.12	.22	.06	.03	-.01	.08	-.03	.56	.43	.37	.01	.06	-.00	-.28	-.05	.01
Pulse				1.00	.17	.17	.03	.03	.17	.18	.06	.08	.05	.12	.02	-.00	.03	.02	-.00	-.04	-.00	-.04	.70	-.13	-.03	.02	-.09	-.09
Systolic Blood Pressure					1.00	.72	.07	.16	.08	.05	.08	.03	.03	.12	.16	.01	.04	-.02	.03	.26	.13	.21	.26	-.02	.04	-.13	-.05	-.02
Diastolic Blood Pressure						1.00	.08	.19	.13	.03	.04	.03	.01	.10	.16	.01	.06	-.00	.01	.24	.13	.19	.21	.01	.04	-.16	-.03	-.01
Arm Skin Folds							1.00	.55	.01	-.02	.09	.07	.16	.10	-.04	.01	-.04	.06	.01	.14	.11	.09	.07	.04	.01	-.11	.00	.06
Back Skin Folds								1.00	.06	-.01	.12	.04	.10	.07	.04	.02	-.01	.04	-.04	.34	.17	.28	.04	.03	-.01	-.19	-.09	-.03
Hematocrit									1.00	.23	.06	.12	.03	.08	-.07	.02	.00	.01	.03	-.01	.03	-.03	.16	-.07	-.01	-.01	-.08	-.08
White Blood Count										1.00	.07	.11	.02	.04	-.12	.02	-.01	-.03	.04	.00	.01	-.00	.21	-.08	-.07	.01	-.05	-.08
Glucose											1.00	.14	.17	.16	-.07	-.01	-.08	-.00	.03	.10	.07	.07	.09	-.04	.04	-.03	-.06	-.03
Cholesterol												1.00	.16	.17	-.16	-.01	-.05	.03	-.02	.04	-.01	.06	.08	-.06	.01	-.03	-.12	-.17
Uric Acid													1.00	.11	.03	-.03	-.09	.04	-.03	.05	.06	.01	.05	-.01	-.01	-.01	.03	.02	-.01
SGPT														1.00	.10	-.01	.03	.03	-.08	.10	.08	.07	.16	-.02	.03	-.06	-.05	-.03
Bilirubin, Total															1.00	-.04	.29	-.05	-.06	.01	-.00	.01	.01	.04	.03	.00	.07	.08
Urine Red Cells																1.00	.10	.02	.02	.01	.00	.00	.00	.01	-.05	-.03	-.05	-.05
Urine White Cells																	1.00	.12	-.03	-.02	-.02	-.00	.03	-.06	.01	-.00	-.01	.01
Specific Gravity																		1.00	-.14	.05	.03	.04	.02	-.04	-.00	-.05	-.03	-.02
pH																			1.00	.03	.05	.00	.02	.02	.04	.02	-.02	-.03
Total Heart Diameter																				1.00	.48	.83	-.07	.10	.06	-.25	-.05	.01
Thoracic Diameter																					1.00	-.08	-.03	.08	.06	-.09	.41	.33
T.D./TH.D.																						1.00	-.05	.06	.04	-.24	-.32	-.21
EKG Heart Rate																							1.00	-.15	-.04	.03	-.06	-.06
PR Interval																								1.00	-.01	-.10	.07	.12
QRS Duration																									1.00	.00	.00	.01
QRS Axis																										1.00	.07	.01
Vital Capacity																											1.00	.76
Forced Expiration Volume																												1.00

either normal or abnormal: a univariate test and a multivariate test.

Univariate Test

The univariate test consists of independently comparing each of the patient's clinical values against the appropriate age group norms. The following hypothesis is tested for each of the patient's values:

H_o: The patient is normal with α being the probability of rejecting H_o when it should be accepted.

Each of the p variables for each of the age groups is normally distributed with a mean and variance. Therefore, the z-test statistic is used

$$z = \frac{x_i - \mu_i}{\sigma_i} \quad . \tag{25}$$

If $|z| \leq z_{\alpha/2}$, H_o is accepted and the patient is classified as normal. On the other hand, if $|z| > z_{\alpha/2}$, then H_o is rejected and the patient is classified as abnormal.

Multivariate Test (11)

It was initially assumed that age and sex were the only independent variables which possibly affect the other variables but are not affected by the other variables. However, height and weight can also be considered independent variables. Therefore, the multivariate normal density function must be made a conditional density function to height and weight.

$\underline{X}$ has been defined as a (p x 1) vector which is normally distributed with mean $\underline{\mu}$ and covariance $\underline{V}$. $\underline{X}$ can be partitioned into two subvectors such that

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}, \tag{26}$$

where $\underline{x}_1$ is the set of clinical variables (x_3, x_4, ..., x_p) and $\underline{x}_2$ is the set of clinical variables (x_1 and x_2).

Also,

$$\underline{x}^* = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2^* \end{bmatrix}, \tag{27}$$

where $\underline{x}_2^*$ is the set of specific values of the patient's clinical variables x_1^* and x_2^*.

The corresponding partitions of the mean vector and the covariance matrix are

$$\underline{\mu} = \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \end{bmatrix}, \text{ and } \qquad \underline{V} = \begin{bmatrix} \underline{V}_{11} & \underline{V}_{12} \\ \underline{V}_{21} & \underline{V}_{22} \end{bmatrix}. \tag{28}$$

Then the conditional distribution $f(\underline{x}_1 | \underline{x}_2^*)$ of the (p-2 x 1) vector $\underline{x}_1$, given the (2 x 1) vector $\underline{x}_2 = \underline{x}_2^*$, is a multivariate normal distribution with mean

$$\underline{u}_1 + \underline{V}_{12} \underline{V}_{22}^{-1} (\underline{x}_2^* - \underline{u}_2), \tag{29}$$

and covariance

$$\underline{V}_{11} - \underline{V}_{12} \underline{V}_{22}^{-1} \underline{V}_{21}. \tag{30}$$

It is possible that values may have not been obtained on all the clinical variables for a patient. Several reasons for not having the clinical values are failure to input the values into the data base, failure of a particular test, or that the particular test was not conducted. In these instances where there are missing values, the (p - 2) vector containing the values of the patient's clinical variables would have some missing values. It should be noted that if the patient's height and weight are missing the test for normality cannot be made.

Since $\underline{x}$ is normally distributed with mean $\underline{\mu}$ and covariance $\underline{V}$, then any subset of $s < p$ is a s-variate normal density function.

Mathematically, if

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}, \tag{31}$$

where $\underline{x}_1$ is a vector containing s of the elements of $\underline{x}$, then $\underline{x}_1$ is a s-variate normal density function. The subset $\underline{x}_1$ could contain the values of those patient's variables which are known, while the subset $\underline{x}_2$ could contain those patient values which are missing.

The probability that $\underline{x}$ lies inside the p-dimensional ellipsoid is given by the inequality

$$(\underline{x} - \underline{\mu})' \underline{V}^{-1} (\underline{x} - \underline{\mu}) \leq c, \tag{32}$$

where c is a known constant.

The quadratic form of the multivariate normal distribution has a chi-square distribution with p degrees of freedom, where p is the rank of the covariance matrix $\underline{V}$. Therefore,

$$(\underline{x} - \underline{\mu})' \underline{V}^{-1} (\underline{x} - \underline{\mu}) \sim \chi^2(p). \tag{33}$$

Now let $\chi^2_{1-\alpha}(p)$ be defined as the upper $100(1 - \alpha)$ percent of the chi-square distribution. Therefore $\chi^2_{1-\alpha}(p)$ can be subsituted for c in the above inequality giving

$$(\underline{x} - \underline{\mu})' \underline{V}^{-1} (\underline{x} - \underline{\mu}) \leq \chi^2_{1-\alpha}(p). \tag{34}$$

Since the conditional distribution $f(\underline{x}_1 | \underline{x}_2^*)$ has been defined, the above inequality would appear as follows when the conditional mean and covariance are substituted

$$\left\{\underline{x}_1 - \left[\underline{u}_1 + \underline{V}_{12}\ \underline{V}_{22}^{-1}\ (\underline{x}_2 - \underline{u}_2)\right]\right\}\quad (\underline{V}_{11} - \underline{V}_{12}\ \underline{V}_{22}^{-1}\ \underline{V}_{21})$$

$$\left\{\underline{x}_1 - \left[\underline{u}_1 + \underline{V}_{12}\ \underline{V}_{22}^{-1}\ (\underline{x}_2 - \underline{u}_2)\right]\right\} \leq \chi^2_{1-\alpha}\,(p-2). \qquad (35)$$

The hypothesis for testing that the patient is normal can now be stated as

H_o: The patient is normal with α being the probability of rejecting H_o when it should be accepted.

If the value of the above inequality is less than or equal to $\chi^2_{1-\alpha}\,(p-2)$, H_o is accepted and the patient is considered normal. On the other hand, if the above inequality is greater than $\chi^2_{1-\alpha}\,(p-2)$, H_o is rejected and the patient is considered abnormal.

Major Contributors to Chi-Square Value

Morrison (16) in the development of the linear discriminant function

$$y = (\underline{x} - \underline{\mu})'\ \underline{V}^{-1}\ \underline{z} \qquad (36)$$

states that if the variances of the variables are nearly

equal, the coefficients

$$b = \underline{V}^{-1} (\underline{x} - \underline{\mu}) \tag{37}$$

give a relative importance of the contribution of each measurement to the χ^2 statistic.

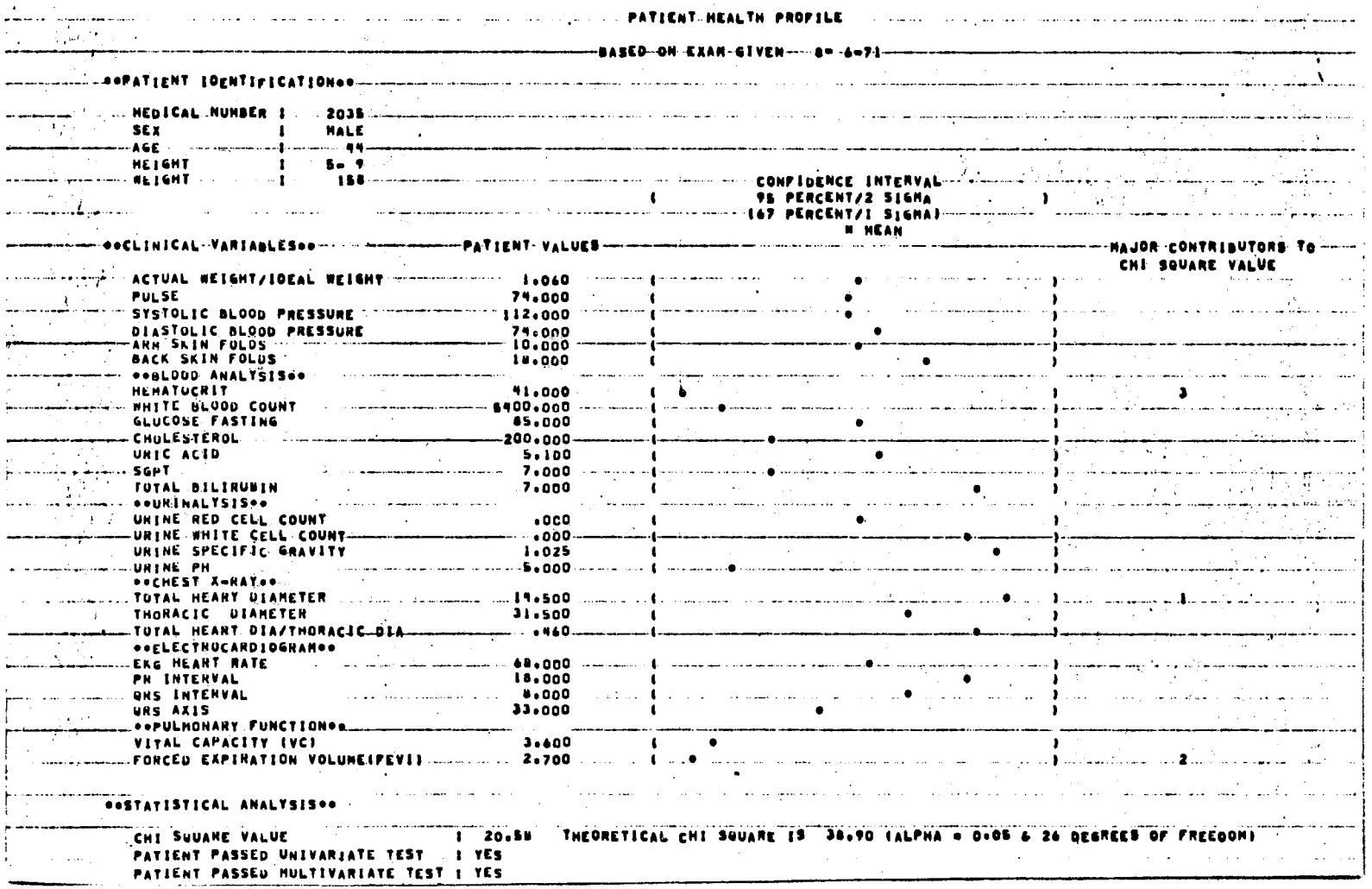
Model Outputs

Three basic outputs have been designed to provide the physician with a useful tool to assist him in analyzing a patient's health. These three outputs are: 1) the patient health profile, 2) the patient longitudinal drift, and 3) the patient summary. Each of these outputs is discussed in the following paragraphs.

Patient Health Profile

The patient health profile is the principle output and is the primary concern of this research. The mathematics associated with the development of a patient's health profile has been thoroughly discussed in the previous sections of this chapter.

Several patient health profiles are presented in Figures 13, 14, and 15. The univariate test consists of checking if any of the twenty-six variables fall outside the 95 percent confidence limits. The multivariate test



PATIENT HEALTH PROFILE

BASED ON EXAM GIVEN 8- 6-71

PATIENT IDENTIFICATION

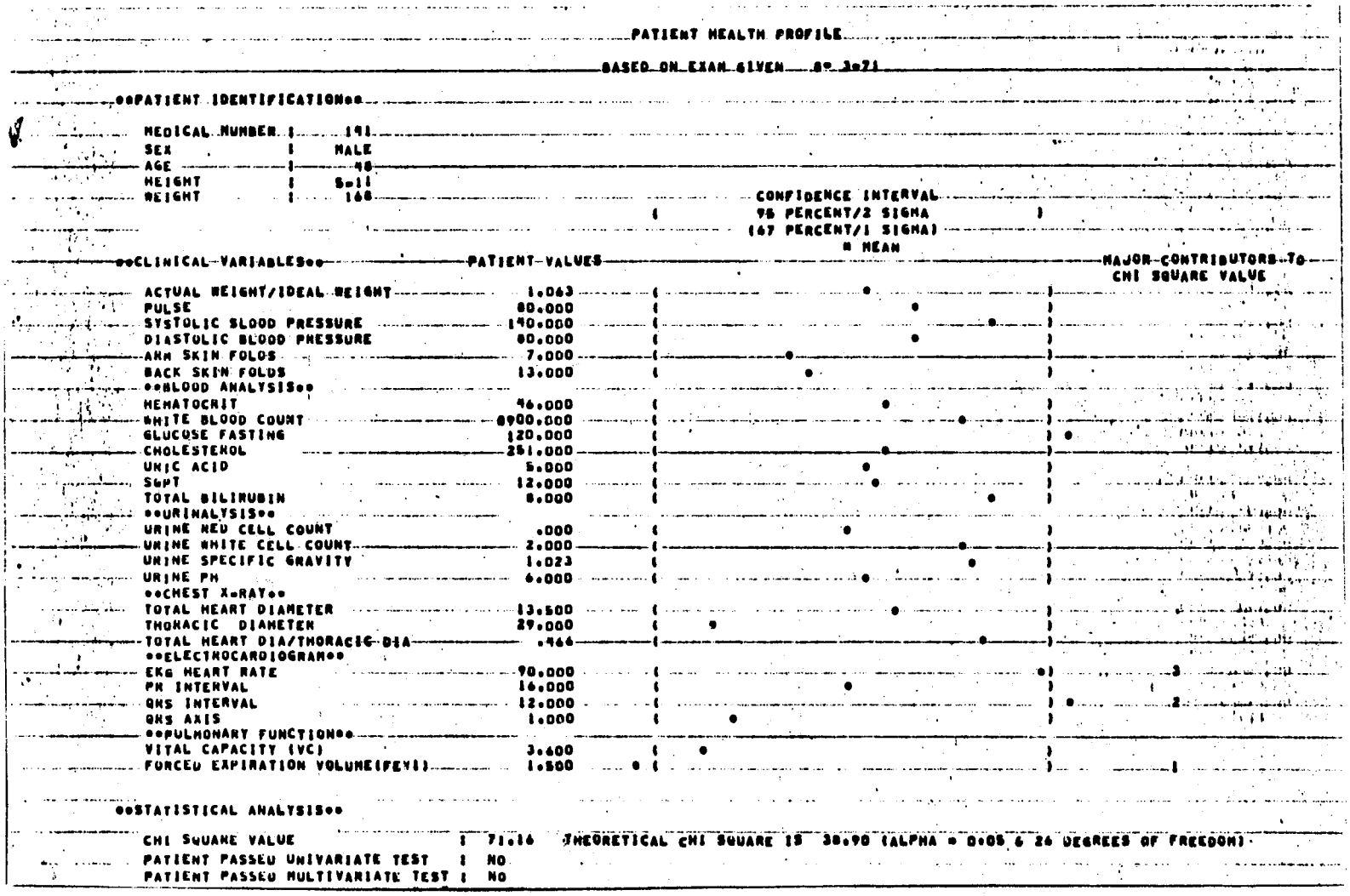
MEDICAL NUMBER	2035
SEX	MALE
AGE	44
HEIGHT	5- 9
WEIGHT	158

CLINICAL VARIABLES	PATIENT VALUES	CONFIDENCE INTERVAL (95 PERCENT/2 SIGMA) (67 PERCENT/1 SIGMA) M MEAN	MAJOR CONTRIBUTORS TO CHI SQUARE VALUE
ACTUAL WEIGHT/IDEAL WEIGHT	1.060		
PULSE	74.000		
SYSTOLIC BLOOD PRESSURE	112.000		
DIASTOLIC BLOOD PRESSURE	74.000		
ARM SKIN FOLDS	10.000		
BACK SKIN FOLDS	18.000		
BLOOD ANALYSIS			
HEMATOCRIT	41.000		3
WHITE BLOOD COUNT	6400.000		
GLUCOSE FASTING	85.000		
CHOLESTEROL	200.000		
URIC ACID	5.100		
SGPT	7.000		
TOTAL BILIRUBIN	7.000		
URINALYSIS			
URINE RED CELL COUNT	.000		
URINE WHITE CELL COUNT	.000		
URINE SPECIFIC GRAVITY	1.025		
URINE PH	5.000		
CHEST X-RAY			
TOTAL HEART DIAMETER	14.500		1
THORACIC DIAMETER	31.500		
TOTAL HEART DIA/THORACIC DIA	.460		
ELECTROCARDIOGRAM			
EKG HEART RATE	68.000		
PR INTERVAL	18.000		
QRS INTERVAL	8.000		
QRS AXIS	33.000		
PULMONARY FUNCTION			
VITAL CAPACITY (VC)	3.600		
FORCED EXPIRATION VOLUME(FEV1)	2.700		2

STATISTICAL ANALYSIS

CHI SQUARE VALUE	20.58	THEORETICAL CHI SQUARE IS 38.90 (ALPHA = 0.05 & 26 DEGREES OF FREEDOM)
PATIENT PASSED UNIVARIATE TEST	YES	
PATIENT PASSED MULTIVARIATE TEST	YES	

Figure 13. Patient Health Profile 1



PATIENT HEALTH PROFILE

BASED ON EXAM GIVEN 8- 3-71

PATIENT IDENTIFICATION

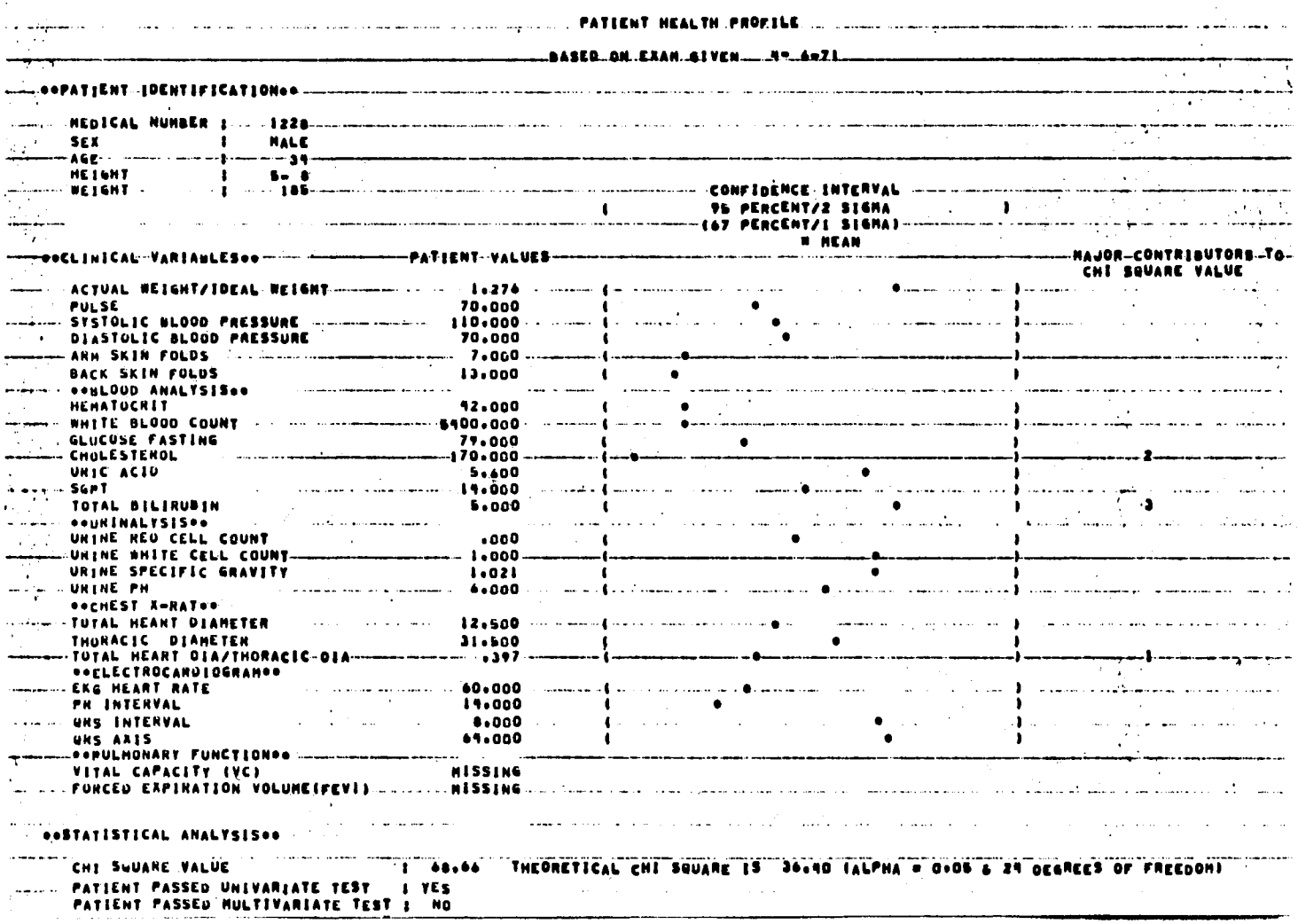
MEDICAL NUMBER	191
SEX	MALE
AGE	48
HEIGHT	5-11
WEIGHT	168

CLINICAL VARIABLES	PATIENT VALUES	CONFIDENCE INTERVAL (95 PERCENT/2 SIGMA) (67 PERCENT/1 SIGMA) * MEAN	MAJOR CONTRIBUTORS TO CHI SQUARE VALUE
ACTUAL WEIGHT/IDEAL WEIGHT	1.063		
PULSE	80.000		
SYSTOLIC BLOOD PRESSURE	140.000		
DIASTOLIC BLOOD PRESSURE	80.000		
ARM SKIN FOLDS	7.000		
BACK SKIN FOLDS	13.000		
BLOOD ANALYSIS			
HEMATOCRIT	46.000		
WHITE BLOOD COUNT	8900.000		
GLUCOSE FASTING	120.000		
CHOLESTEROL	261.000		
URIC ACID	5.000		
SGPT	12.000		
TOTAL BILIRUBIN	8.000		
URINALYSIS			
URINE RED CELL COUNT	.000		
URINE WHITE CELL COUNT	2.000		
URINE SPECIFIC GRAVITY	1.023		
URINE PH	6.000		
CHEST X-RAY			
TOTAL HEART DIAMETER	13.500		
THORACIC DIAMETER	29.000		
TOTAL HEART DIA/THORACIC DIA	.466		
ELECTROCARDIOGRAM			
EKG HEART RATE	90.000		3
PR INTERVAL	16.000		
QRS INTERVAL	12.000		2
QRS AXIS	1.000		
PULMONARY FUNCTION			
VITAL CAPACITY (VC)	3.600		
FORCED EXPIRATION VOLUME(FEV1)	1.500		1

STATISTICAL ANALYSIS

CHI SQUARE VALUE	71.16	THEORETICAL CHI SQUARE IS 38.98 (ALPHA = 0.05 & 24 DEGREES OF FREEDOM)
PATIENT PASSED UNIVARIATE TEST	NO	
PATIENT PASSED MULTIVARIATE TEST	NO	

Figure 14. Patient Health Profile 2



PATIENT HEALTH PROFILE

BASED ON EXAM GIVEN N= 4871

PATIENT IDENTIFICATION

MEDICAL NUMBER	1228
SEX	MALE
AGE	34
HEIGHT	5- 8
WEIGHT	185

CLINICAL VARIABLES	PATIENT VALUES	CONFIDENCE INTERVAL (95 PERCENT/2 SIGMA) (67 PERCENT/1 SIGMA) * MEAN	MAJOR CONTRIBUTORS TO CHI SQUARE VALUE
ACTUAL WEIGHT/IDEAL WEIGHT	1.276		
PULSE	70.000		
SYSTOLIC BLOOD PRESSURE	110.000		
DIASTOLIC BLOOD PRESSURE	70.000		
ARM SKIN FOLDS	7.000		
BACK SKIN FOLDS	13.000		
BLOOD ANALYSIS			
HEMATOCRIT	42.000		
WHITE BLOOD COUNT	6400.000		
GLUCOSE FASTING	79.000		
CHOLESTEROL	170.000		2
UNIC ACID	5.600		
SGPT	19.000		
TOTAL BILIRUBIN	5.000		3
URINALYSIS			
URINE RED CELL COUNT	.000		
URINE WHITE CELL COUNT	1.000		
URINE SPECIFIC GRAVITY	1.021		
URINE PH	6.000		
CHEST X-RAY			
TOTAL HEART DIAMETER	12.500		
THORACIC DIAMETER	31.500		
TOTAL HEART DIA/THORACIC DIA	.397		1
ELECTROCARDIOGRAM			
EKG HEART RATE	60.000		
PR INTERVAL	14.000		
QRS INTERVAL	8.000		
QRS AXIS	64.000		
PULMONARY FUNCTION			
VITAL CAPACITY (VC)	MISSING		
FORCED EXPIRATION VOLUME(FEV1)	MISSING		

STATISTICAL ANALYSIS

CHI SQUARE VALUE	68.66	THEORETICAL CHI SQUARE IS 36.40 (ALPHA = 0.05 & 24 DEGREES OF FREEDOM)
PATIENT PASSED UNIVARIATE TEST	YES	
PATIENT PASSED MULTIVARIATE TEST	NO	

Figure 15. Patient Health Profile 3

involves comparing the patient's actual chi-square value with the theoretical chi-square value for $\alpha = 0.05$ and the appropriate degrees of freedom.

The patient's profile in Figure 13 is classified normal since he passed both the univariate and multivariate tests. His chi-square value is 20.58 which is less than the theoretical chi-square value of 38.90. The major contributors to the patient's chi-square value are his high total heart diameter, low FEV_1, and low hematocrit.

The patient's profile in Figure 14 is definitely abnormal since he failed both the univariate and multivariate tests. His chi-square value is 71.16 which is greater than the theoretical chi-square value of 38.90. The patient's glucose and QRS duration are greater than the two sigma limits and his FEV_1 is less than the two sigma limits. The major contributors to his chi-square value are a low FEV_1, high QRS duration, and high EKG heart rate.

The patient's profile in Figure 15 is more interesting. This patient passed the univariate test; however, because of the correlation pattern existing between his variables, failed the multivariate test. It is in these instances that, even though the patient's values are within limits, there is something that may be severly abnormal. Looking at the major contributors to the chi-square value indicates

that the patient has a low T.D./TH.D. ratio, a low cholesterol, and a high bilirubin. This additional information should provide the physician with a starting point in diagnosing the patient.

Patient Longitudinal Drift

An output of the patient health profile is his chi-square value. This value can be considered as a composite measure of the patient's health which takes into consideration all the variables and the correlation pattern between the variables. Therefore, this single value could be used to trace the patient's drift over time; that is, from exam to exam. A typical plot of this composite value over time is given in Figure 16.

Patient Summary

As an additional tool to assist the physician in analyzing a patient's health, it is desirable that the physician have some type of summary of the patient's medical file. One such type of patient summary, using the Medical Center's data base, is the six page report in Tables XV through XX.

The first page (Table XV) of the report contains the patient's personal data. This data is from the history form

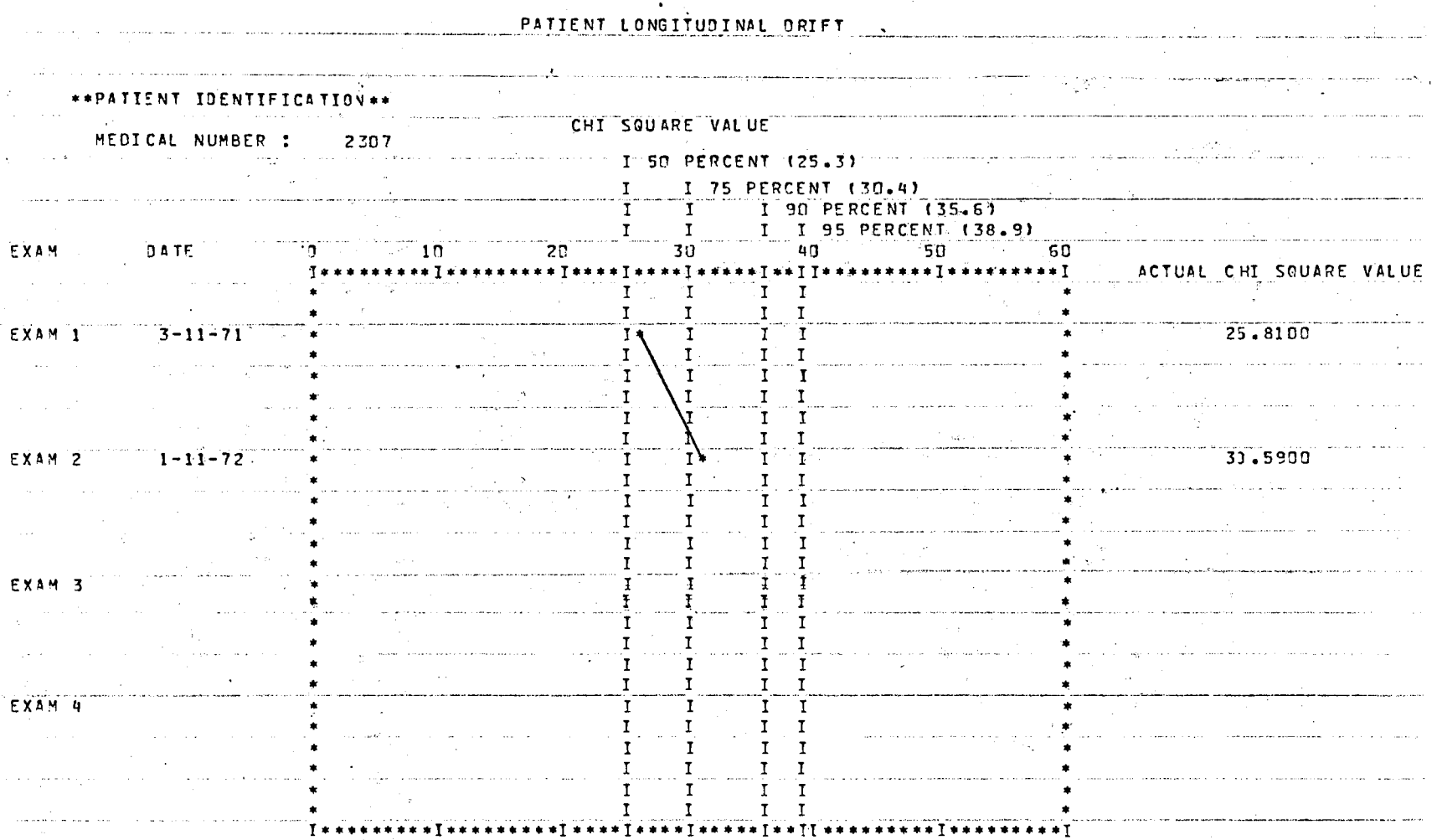


Figure 16. Patient Longitudinal Drift

TABLE XV

PATIENT SUMMARY-PERSONAL DATA

NASA/MSFC MEDICAL AUTOMATION SYSTEM - PATIENT SUMMARY REPORT

HISTORY DATA

784 : PATIENT MEDICAL NUMBER

EMPLOYEE DATA

SEX	: M		
BIRTH DATE MO/DA/YR	: 121726		
PERFORMING ACTIVITY	: 340		
OCCUPATION CODE	: 13010		
NASA CLASS CODE	: 70000		
GS/WB LEVEL	: PL00		
AGE	: 49.		

PRESENT MEDICATIONS (1=YES 0=NO) FOR

BLOOD PRESSURE	: 0
ANTI-ANGINAL	: 0
DIABETES	: 0
ARTHRITIS	: 0
HORMONES	: 0
TRANQUILIZERS	: 0
OTHER	: 0

EMPLOYEE STATUS

STATUS CODE	: 1
SEPARATION CODE	: 0
SEPARATION DATE	: 0
REASON FOR SEPARATION IF DUE TO HEALTH	
DIAGONIS CODE	:
DIAGONIS CODE	:
DIAGONIS CODE	:

SMOKING HABIT (1=YES 2=NO)

EVER SMOKED PIPE OR CIGAR	: 2
EVER SMOKED CIGARETTES	: 2
YEARS SMOKED	: 0
PRESENTLY SMOKING CIGARETTES	: 2
AVG CIGARETTES A DAY	: 0

FAMILY DISEASE HISTORY (1=YES 2=NO)

OF DIABETES	: 0
OF HEART DISEASE	: 0

PREVIOUS MEDICAL CONDITIONS RELATING TO

DENIAL OF LIFE INSURANCE	DIAGNOSIS CODE/DATE :
OPERATIONS	DIAGNOSIS CODE/DATE :
HOSPITALIZATION	DIAGNOSIS CODE/DATE :
OTHER INJURY OR ILLNESS	DIAGNOSIS CODE/DATE :
TREATMENT WITHIN PAST 5 YEARS	DIAGNOSIS CODE/DATE :
REJECTION FOR MILITARY SERVICE	DIAGNOSIS CODE/DATE :
MILITARY DISCHARGE BECAUSE OF HEALTH	DIAGNOSIS CODE/DATE :

TABLE XVI

PATIENT SUMMARY-PRIOR ABNORMALITIES

NASA/MSFC MEDICAL AUTOMATION SYSTEM - PATIENT SUMMARY REPORT

PHYSICIAL EXAM DIAGNOSES

784 : PATIENT MEDICAL NUMBER

PATIENT DIAGNOSES FROM EXAM 1 GIVEN 7-11-68

429.9	ILL-DEFINED HEART DISEASE
744.3	CONGENITAL ANOMALIES OF EYE
569.1	OTHER DISEASES OF INTESTINES AND PERITONEUM

PATIENT DIAGNOSES FROM EXAM 2 GIVEN 4-18-69

744.3	CONGENITAL ANOMALIES OF EYE

PATIENT DIAGNOSES FROM EXAM 3 GIVEN 5-26-70

725.1	DISPLACEMENT OF INTERVERTEBRAL DISC

PATIENT DIAGNOSES FROM EXAM 4 GIVEN 5-14-71

725.1	DISPLACEMENT OF INTERVERTEBRAL DISC
744.3	CONGENITAL ANOMALIES OF EYE

TABLE XVII

PATIENT SUMMARY-PHYSICAL EXAMINATION

NASA/MSFC MEDICAL AUTOMATION SYSTEM - PATIENT SUMMARY REPORT

PHYSICAL EXAM DATA

784 : PATIENT MEDICAL NUMBER	EXAM 1	EXAM 2	EXAM 3	EXAM 4	EXAM 5	EXAM 6	EXAM 7	EXAM 8	EXAM 9	EXAM 10
EXAM DATE YR/MO/DA :	680711	690418	700526	710514						
EXAM TYPE (1=PREPLAC 2=RE-EMP 3=PERIODIC 6=EXEC 7=MONITOR) :	3	3	3	6						
WORK CLASSIFICATION :	0	1	0	1						
EXAMINING PHYSICIAN :	0	0	0	2						
PHYSICAL CHARACTERISTICS										
HEIGHT :	5-10	5-10	5-10	5-11						
WEIGHT :	207.	186.	193.	195.						
IDEAL WEIGHT :	0.	0.	0.	158.						
PULSE :	84.	70.	60.	84.						
SYSTOLIC BLOOD PRESSURE :	104.	110.	90.	102.						
DIASTOLIC BLOOD PRESSURE :	75.	68.	60.	70.						
RECUMBENT SYS BLOOD PRESS :	0.	0.	0.	0.						
RECUMBENT DIA BLOOD PRESS :	0.	0.	0.	0.						
ARM SKIN FOLDS :	0.	0.	0.	6.						
BACK SKIN FOLDS :	0.	0.	0.	24.						
VISION										
COLOR VISION :										
FAR SIGHT R EYE UNAIDED :	0.	0.	0.	0.						
FAR SIGHT R EYE CORRECTED :	0.	100.	30.	30.						
FAR SIGHT L EYE UNAIDED :	0.	0.	0.	20.						
FAR SIGHT L EYE CORRECTED :	20.	30.	20.	20.						
NEAR SIGHT R EYE UNAIDED :	0.	0.	100.	200.						
NEAR SIGHT R EYE CORRECTED :	20.	20.	0.	0.						
NEAR SIGHT L EYE UNAIDED :	0.	0.	0.	100.						
NEAR SIGHT L EYE CORRECTED :	0.	100.	20.	0.						
PROCTOSCOPIC (1=NEG 2=BENIGN POLYP 3=CANCER 4=OTHER) :	0	0	0	1						
DIAGNOSIS (STATUS CODE=1 NEW DIAGNOSIS CODE=0 SAME DIAGNOSIS)										
DIAGNOSIS CODE 1 :	429.9	744.3	725.1	725.1						
STATUS CODE 1 :	0.	0.	0.	0.						
DIAGNOSIS CODE 2 :	744.3	.0	.0	744.3						
STATUS CODE 2 :	0.	0.	0.	0.						
DIAGNOSIS CODE 3 :	569.1	.0	.0	.0						
STATUS CODE 3 :	0.	0.	0.	0.						
DIAGNOSIS CODE 4 :	.0	.0	.0	.0						
STATUS CODE 4 :	0.	0.	0.	0.						
DIAGNOSIS CODE 5 :	.0	.0	.0	.0						
STATUS CODE 5 :	0.	0.	0.	0.						

TABLE XVIII

PATIENT SUMMARY-DOCTOR'S COMMENTS

NASA/MSFC MEDICAL AUTOMATION SYSTEM - PATIENT SUMMARY REPORT

PHYSICAL EXAM DATA CONTINUED

784 : PATIENT MEDICAL NUMBER	EXAM 1	EXAM 2	EXAM 3	EXAM 4	EXAM 5	EXAM 6	EXAM 7	EXAM 8	EXAM 9	EXAM 10
GENERAL CHARACTERISTICS (0=NORMAL 1=ABNORMAL 2=NOT EXAMINED)										
GENERAL APPEARANCE :	0	0	0	0						
POSTURE :	0	0	0							
GAIT :	0	0	0	0						
BEHAVIOR :	0	0	0	0						
HEAD :	0	0	0	0						
EYES :	1	1	1	1						
FUNDI :	1	0	0	1						
EARS :	0	0	0	0						
NMT :	0	0	0	1						
NECK :	0	0	0	0						
UPPER EXTREMITY :	0	0	0	0						
THORAX :	0	0	0	0						
BREASTS :	0	0	0	0						
LUNGS :	0	0	0	0						
HEART :	1	0	0	0						
ABDOMENT :	0	0	0	0						
GROIN-GENITALIA :	0	0	0	0						
BACK :	0	0	1	0						
PELVIS :	0	0	0	0						
RECTUM-ANUS :	0	0	0	0						
PROSTATE :	0	0	0	0						
FEMALE PELVIS :	2	2	2	2						
LOWER EXTREMITY :	0	0	0	0						
ARTERIES AND VEINS :	0	0	0	0						
JOINTS :	0	0	0	0						
MUSCLES :	0	0	0	0						
NEUROLOGICAL :	0	0	1	1						
SKIN :	0	0	0	0						
LYMPH NODES :	0	0	0	0						
OTHER :	0	0	0	0						
X-RAYS (1=NORMAL 2=SUSPICIOUS 3=ABNORMAL)										
CHEST X-RAY :	1	0	1	1						
BACK X-RAY :	0	0	0	0						
KUB X-RAY :	1	0	1	0						
OTHER X-RAY :	0	0	0	0						

TABLE XIX

PATIENT SUMMARY-LABORATORY RESULTS

NASA/MSFC MEDICAL AUTOMATION SYSTEM - PATIENT SUMMARY REPORT

LABORATORY DATA

784 : PATIENT MEDICAL NUMBER	EXAM 1	EXAM 2	EXAM 3	EXAM 4	EXAM 5	EXAM 6	EXAM 7	EXAM 8	EXAM 9	EXAM 10
LABORATORY DATE YR/MO/DA	680703	690921	700520	710511						
HEMATOCRIT	44.	42.	43.	45.						
PLATELETS	N	N	N	N						
WHITE BLOOD COUNT	68.	76.	71.	61.						
GLUCOSE FASTING	60.	123.	98.	91.						
GLUCOSE TWO HOUR	0.	0.	0.	0.						
CHOLESTEROL	195.	205.	240.	210.						
URIC ACID	4.2	5.6	6.3	6.1						
SGPT	10.	20.	8.	10.						
THYMOL TURBIDITY	0.	0.	0.	0.						
BILIRUBIN DIRECT	0.	0.	0.	0.						
BILIRUBIN INDIRECT	2.	6.	4.	0.						
BILIRUBIN TOTAL	2.	6.	4.	4.						
URINALYSIS										
RED CELLS	0.	0.	0.	0.						
WHITE CELLS	0.	0.	0.	3.						
SPECIFIC GRAVITY	.000	1.019	1.019	1.015						
SUGAR	0.	0.	0.	0.						
ALBUMIN	0.	0.	0.	0.						
PH	0.	6.	6.	6.						
CHEST X-RAY										
TOTAL HEART DIAMETER	0.	0.	0.	120.						
THORACIC DIAMETER	0.	0.	0.	320.						
HEART SIZE	0.	0.	0.	1.						
AORTA	0.	0.	0.	1.						
PARENCHYMAL	0.	0.	0.	8.						
ELECTROCARDIOGRAM										
HEART RATE	72.	74.	66.	62.						
P-R INTERVAL	0.	0.	0.	18.						
QRS COMPLEX	0.	0.	0.	8.						
QRS AXIS	35.	40.	30.	30.						
NORMAL/ABNORMAL				N						
ABNORMAL CONDITIONS										
Q-WAVE	0.	0.	0.	0.						
R-WAVE	0.	0.	0.	0.						
S-T SEGMENT	0.	0.	0.	0.						
T-WAVE	0.	0.	0.	0.						
A-V CONDUCTION	0.	0.	0.	0.						
VENTRIC CONDUCTION	0.	0.	0.	0.						
ARRHYTHMIAS	0.	0.	0.	0.						
MISCELLANEOUS	0.	0.	0.	0.						
PULMONARY FUNCTION										
VITAL CAPACITY	.0	.0	.0	5.3						
FORCED EXPIRATION VOLUME	.0	.0	.0	3.9						

TABLE XX

PATIENT SUMMARY-OCCUPATIONAL STATISTICS

" NASA/MSFC MEDICAL AUTOMATION SYSTEM - PATIENT SUMMARY REPORT

NON SCHEDULED VISITS

798 : PATIENT MED NUMBER	VISIT 1	VISIT 2	VISIT 3	VISIT 4	VISIT 5	VISIT 6	VISIT 7	VISIT 8	VISIT 9	VISIT10	VISIT11	VISIT12
SEX	M	M										
AGE	44.	44.										
ORGANIZATION CODE	0.	340.										
TREATMENT CODE	31.	0.										
DATE OF VISIT	591009.	710511.										
TYPE OF EXAMINATION	0.	4.										
TYPE OF PROCEDURE		OTHER										
TYPE OF LAB TEST		HMTLGY										
HEALTH COUNSELING (1=INDIV 2=GROUP)	0.	0.										
IMMUNIZATION (1=YES 2=NO)	0.	0.										
DISPOSITION	0.	1.										
PROFESSIONAL EFFORT	0.	4.										
SOURCE OF VISIT	0.	3.										
FOLLOWUP CODE	0.	0.										
NUMBER OF DAYS LOST	0.	0.										
DIAGNOSIS												
DIAGNOSIS CODE	7179	7251										
CONDITION (0=OLD 1=NEW)	0.	0.										
FOLLOWUP OUTCOME	0.	0.										
DIAGNOSIS CODE	7296	7443										
CONDITION (0=OLD 1=NEW)	0.	0.										
FOLLOWUP OUTCOME	0.	0.										
DIAGNOSIS CODE												
CONDITION (0=OLD 1=NEW)	0.	0.										
FOLLOWUP OUTCOME	0.	0.										

which is completed at the patient's first visit and thereafter updated when necessary. The patient history data includes job status, any present medications, smoking habits, any family history of diabetes and heart disease, and a list of diagnosis codes describing his previous health.

The second page (Table XVI) of the report summarizes the patient's prior abnormalities as diagnosed by the physician. For this particular summary, the patient has had four physicals. On his first exam he was diagnosed as having an ill-defined heart disease, a congenital anomaly of the eye, and an other disease of the intestines and peritoneum. On his second exam he was diagnosed as having a congenital anomaly of the eye. On his third exam he had a displacement of an intervertebral disc. His fourth exam indicated a displacement of the intervertebral disc and a congenital anomaly of the eye.

The third and fourth pages of the patient summary present the detailed results of the physical examination. Included on the third page (Table XVII) are the patient's physical characteristics, vision, proctoscopic findings, and the physician's diagnosis. The fourth page (Table XVIII) contains the doctor's opinion of the general characteristics of the patient and the summary x-ray results.

The fifth (Table XIX) presents the patient's detailed laboratory results. It is from the lab results that the majority of the values are taken for determining a patient's health profile. Included on the fifth page are the results of the blood analysis, urine analysis, chest x-ray, electrocardiogram, and pulmonary functions.

The sixth and last page (Table XX) of the patient summary contains the data from the occupational statistics form. This page is a summary of all the patient's visits, both scheduled and nonscheduled, to the Medical Center.

Computer Programming and Hardware Requirements

A set of five computer programs have been written to define the age group norms. These programs make extensive use of existing statistical routines. All the routines are written in FORTRAN for the Univac 1108 computer. Maximum core requirements are 45K words (36 bits per word). The maximum number of tape drives is two for input and one for output.

In addition to the programs for defining the age group norms, three programs have been written to output the results of a patient analysis. These three programs correspond to the three outputs presented in Figures 13 through 16 and Tables XV through XX. A generalized flow diagram

of the computer requirements for using these three programs is given in Figure 17. All the programs are written in FORTRAN for the Univac 1108. Each of these programs is briefly discussed in the following paragraphs.

The program for computing the health profiles requires 40K of core. The core can be reduced to 25K if single precision is used. Two tapes are required as input. Compilation time for the program is 45 seconds. A patient health profile can be computed in ten seconds.

The program for plotting the patient longitudinal drift requires 7K of core. Compilation time is two seconds. Execution time per plot is one second.

The patient summary is created directly from the medical data base. Two tapes are required for input. Compilation time is 20 seconds. Execution time is six seconds per patient summary.

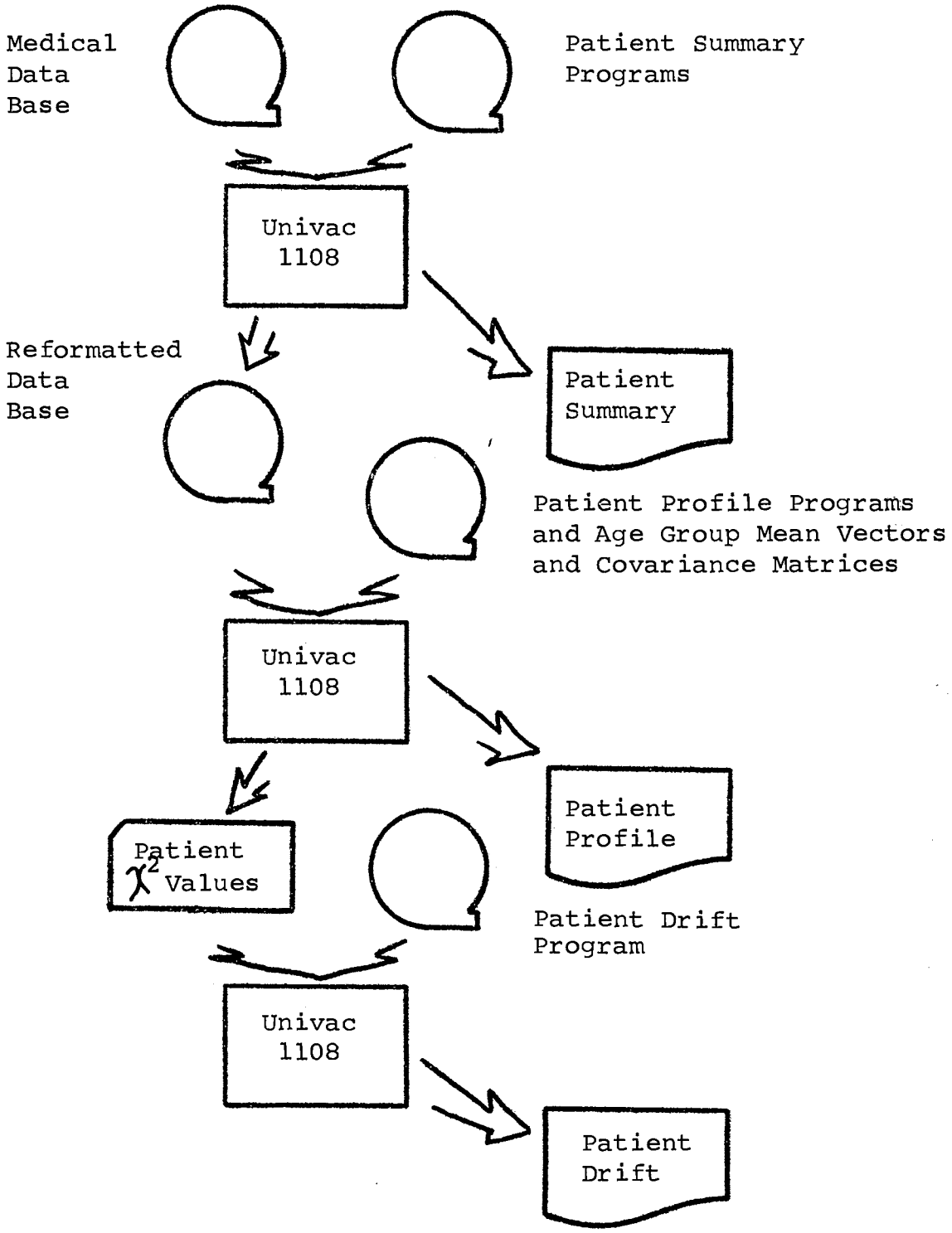


Figure 17. Generalized Flow Diagram of Programs

CHAPTER IV

TEST AND EVALUATION

Chapter IV presents the testing and evaluation of the model. Included in this chapter are the test procedures, the analysis using the original sample, and the analysis using the independent sample.

Test Procedure

The testing of the model consists of two steps. The first step is the taking of a random sample from the data originally used to develop the age group norms. The patients selected from this sample are then compared with the defined age group norms and classified as either normal or abnormal.

The second step in testing the model consists of taking a random sample from an independent data source. These patients are also compared with the age group norms and classified as either normal or abnormal.

Analysis Using Original Sample

The first step in the analysis consists of taking a random sample of 923 patients from the population of 3825 patients. Ninety-five percent confidence limits are established for classifying a patient as normal or abnormal. That is, there is a five percent chance of classifying a patient abnormal when he is actually normal (i.e., a false positive classification). This false positive must be balanced against classifying a patient normal when he is actually abnormal. This is a more serious error and is commonly referred to as a false negative classification.

From a medical point of view, a false negative classification is of more concern. The risk of a false negative can be decreased at the expense of increasing the number of false positives. However, from an economical point of view, it may be infeasible and impractical to decrease the number of false negatives by increasing the risk of false positives. Therefore, there must be a tradeoff between the two.

Model's Classification

The results of the model's classification is given in Table XXI. From the table 58.3 percent of the patients have one or more clinical value falling outside the 95 percent

confidence limits. On the other hand, 16.9 percent of the patients have a chi-square value greater than the theoretical chi-square for a level of significance of $\alpha = 0.05$ and the appropriate degrees of freedom. The number of degrees of freedom vary with each patient depending on the number of missing values.

TABLE XXI

RESULTS OF MODEL'S CLASSIFICATION USING ORIGINAL SAMPLE

		95 Percent Multivariate Chi-square Test		
		Normal	Abnormal	Total
95 Percent Univariate Confidence Limits	Within	374/40.5%	11/ 1.2%	385/ 41.7%
	Outside	393/42.6%	145/15.7%	538/ 58.3%
	Total	767/83.1%	156/16.9%	923/100.0%

Notice that 1.2 percent have no values falling outside the 95 percent univariate confidence limits; but, because of the correlations between the variables, have a chi-square value greater than the theoretical value (i.e., pass the univariate test but fail the multivariate test). Likewise,

42.6 percent have one or more value falling outside the 95 percent confidence limits, but have a chi-square value less than the theoretical value (i.e., fail the univariate test but pass the multivariate test).

The cumulative distribution of the chi-square values for those patients having values recorded for all twenty-eight variables is given in Figure 18. From this figure it can be seen how the number of patients classified as abnormal increase as alpha is increased. For example, those patients who fall between $\alpha = 0.05$ and $\alpha = 0.25$ could be classified as borderline cases which may be classified as abnormal at the time of their next physical. It is for these patients who are borderline cases that a plot of the patient's chi-square values from exam to exam could be helpful in possibly detecting a trend toward abnormality before it actually occurs. Such a plot is shown in Figure 16 in Chapter III.

Model's Classification Versus Doctor's Diagnosis

In order to validate the results of the model's classification, the results of the examining medical doctor's diagnosis are used to compare the model's classification with the doctor's diagnosis. At this point it is assumed that if the patient's medical record for a specific exam

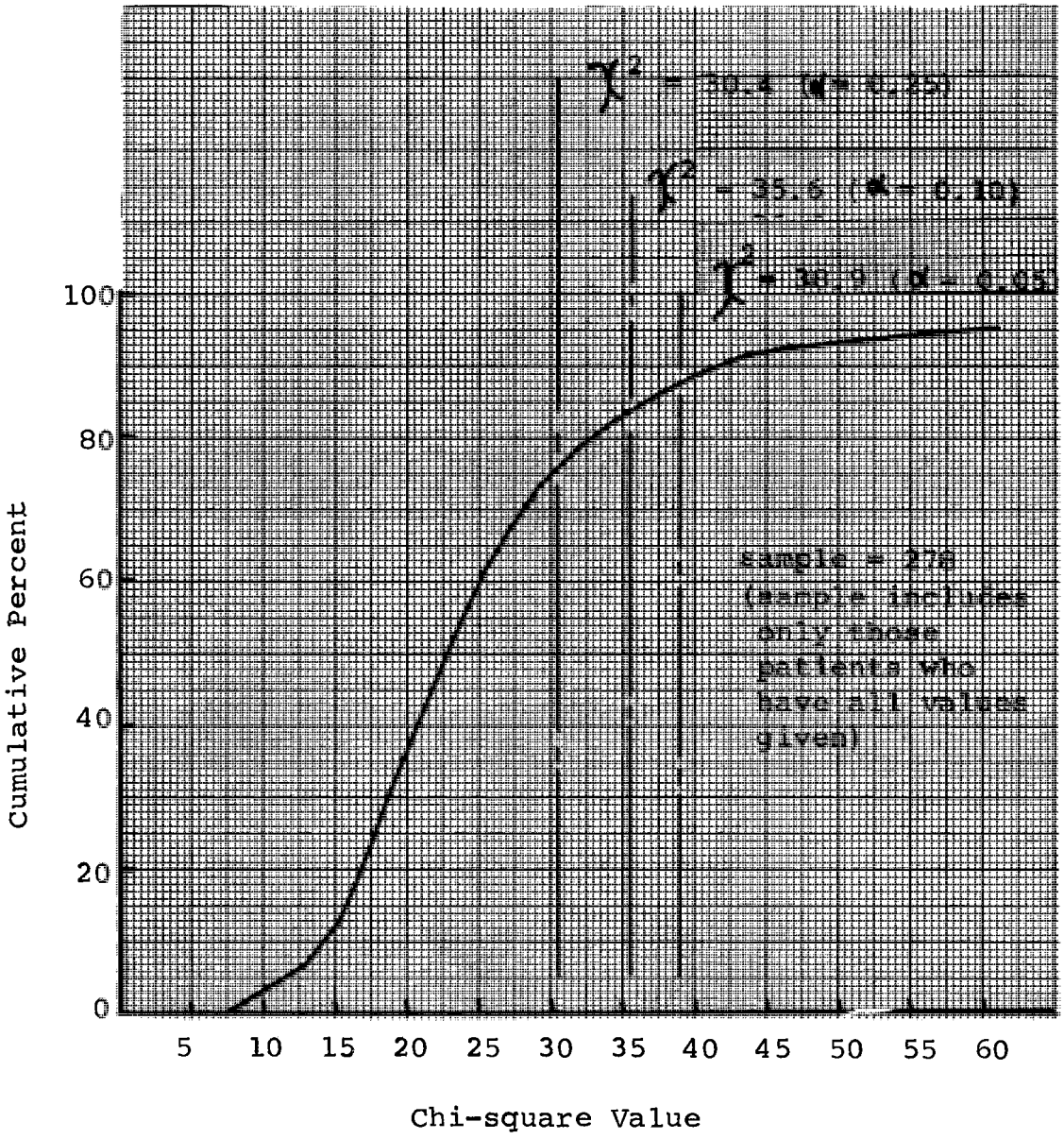


Figure 18. Distribution of Chi-Square Values Using Initial Sample

contains one or more diagnosis codes, the patient is abnormal, regardless of the specific diagnosis. Likewise, a patient with no diagnosis codes is assumed normal.

It is further assumed that if a patient fails either the model's univariate or the multivariate test (or both), the patient is abnormal. Likewise, if a patient passes both the univariate and the multivariate tests he is normal.

The results of the comparison is given in Table XXII. From the table 26.8 percent of the patients who are diagnosed as normal by the doctor are also classified as normal by the model. Likewise, 40.1 percent of the patients who are diagnosed as abnormal by the doctor are also diagnosed as abnormal by the model. This is equivalent of saying, that for 66.9 percent of the sample, the model agreed with the doctor, and for 33.1 percent the model disagreed with the doctor.

To test for any agreement between the model's results and the doctor's diagnosis, the following hypothesis is defined

H_o: The diagnosis resulting from the model's classification is independent of the doctor's diagnosis.

Or,

$$H_o: \quad p_{ij} = p_i \, p_j, \quad i = 1,2, \quad j = 1,2.$$

TABLE XXII

MODEL'S CLASSIFICATION COMPARED WITH DOCTOR'S DIAGNOSIS (USING ORIGINAL SAMPLE)

		Doctor's Diagnosis		
		Normal	Abnormal	Total
Model's Classification	Normal	247/26.8%	127/13.7%	374/ 40.5%
	Abnormal	179/19.4%	370/40.1%	549/ 59.5%
	Total	426/46.2%	497/53.8%	923/100.0%

To test H_o the chi-square test (18) is computed as

$$\chi^2_{(2-1)\ (2-1)} = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(x_{ij} - n\,\hat{p}_{i.}\,\hat{p}_{.j})^2}{n\,\hat{p}_{i.}\,\hat{p}_{.j}}, \qquad (38)$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(x_{ij} - T_{i.}\,T_{.j}/n)^2}{T_{i.}\,T_{.j}/n}, \qquad (39)$$

where x_{ij} = the observed frequencies,

$\hat{p}$'s = the expected proportions, and

T's = the expected frequencies.

Using the data in Table XXII a chi-square value of 99.2 is computed which is greater than the theoretical chi-square value of 3.84 for $\alpha = 0.05$ and one degree of freedom. Therefore, H_o is rejected and there is a dependence between the model's classification and the doctor's diagnosis.

Although there is a dependence between the doctor's diagnosis and the model's classification, for 33.1 percent of the patients there is a disagreement between the doctor's diagnosis and the model's classification. The model classified 19.4 percent of the sample as abnormal while the doctor diagnosed the patients normal. A more serious disagreement is that the model classified 13.7 percent of the sample normal while the doctor diagnosed them abnormal.

Model's Classification Versus Patient Diagnosis

It is of value to further study the 923 patients to determine the prevalence of diseases among this sample. By referring to the patient medical records it is possible to tabulate, based on the doctor's diagnosis, the prevalence of various diseases. The seventeen major classifications of diseases, as defined by the U.S. Department of Health, Education, and Welfare (19) are used rather than the entire list of one thousand diagnoses. The distribution of diseases for the 923 patients is summarized in Table XXIII.

TABLE XXIII

PREVALENCE OF DISEASES FOR ORIGINAL SAMPLE

	Major Disease Group	Prevalence (percent)
000-136	Infective and parasitic diseases	1.8
140-239	Neoplasms	2.3
240-279	Endocrine, nutritional, and metabolic diseases	42.1
280-289	Diseases of the blood and blood forming organs	0.1
290-315	Mental disorders	1.4
320-389	Diseases of the nervous system and sense organs	15.8
390-458	Diseases of the circulatory system	23.2
460-519	Diseases of the respiratory system	4.9
520-577	Diseases of the digestive system	9.3
580-629	Diseases of the genitourinary system	6.4
630-678	Complications of pregnancy, childbirth, and the puerperium	0.1
680-709	Diseases of the skin and subcutaneous tissue	3.1
710-738	Diseases of the musculoskeletal system and connective tissue	10.1
740-759	Congenital anomalies	1.5
760-779	Certain causes of perinatal morbidity and mortality	0.5
780-796	Symptoms and illdefined conditions	1.5
800-999	Accidents, poisonings, and violence	2.4

sample size 923

It is of interest to determine what percentage of those patients, who the doctor diagnosed as having a particular disease, are also classified as abnormal by the model. This should give an indication as to the accuracy of the model in detecting particular diseases. This comparison is given in Table XXIV.

It should be noted that the number of diagnoses does not agree with the sample size. This is because a patient may have more than one diagnosis. The data in Table XXIV only represents the doctor's diagnosis. If the doctor failed to identify a disease, then the patient is considered normal as far as having that disease.

Correction for Nondetectable Diseases

During the comparison of the model's classification with the doctor's diagnosis, it was noticed that the diagnosis for obesity (code 277) varied considerably among physicians. Each physician's diagnosis of obesity is summarized in Table XXV. After discussion with the medical staff, it was decided to ignore the obesity diagnosis in comparing the model's classification with the doctor's diagnosis.

TABLE XXIV

COMPARISON OF MODEL'S CLASSIFICATION VERSUS PATIENT DIAGNOSIS

	Major Disease Group	Number of Patients		Total
		Normal by Model and Abnormal by Doctor	Abnormal by Both Model and Doctor	
000-136	Infective and parasitic diseases	8/ 47.0%	9/ 53.0%	17/100.0%
140-239	Neoplasms	7/ 33.3%	14/ 66.7%	21/100.0%
240-279	Endocrine, nutritional, and metabolic diseases	131/ 33.8%	257/ 66.2%	388/100.0%
280-289	Diseases of the blood and blood forming organs	0/ 0.0%	1/100.0%	1/100.0%
290-315	Mental disorders	1/ 7.6%	12/ 92.4%	13/100.0%
320-389	Diseases of the nervous system and sense organs	46/ 31.5%	100/ 68.5%	146/100.0%
390-458	Diseases of the circulatory system	52/ 24.2%	162/ 75.8%	214/100.0%
460-519	Diseases of the respiratory system	18/ 40.0%	27/ 60.0%	45/100.0%
520-577	Diseases of the digestive system	32/ 37.2%	54/ 62.8%	86/100.0%
580-629	Diseases of the genitourinary system	18/ 30.5%	41/ 69.5%	59/100.0%
630-678	Complications of pregnancy, childbirth, and the puerperium	1/100.0%	0/ 0.0%	1/100.0%
	subcutaneous tissue	14/ 48.2%	15/ 51.8%	29/100.0%
710-738	Diseases of the musculoskeletal system and connective tissue	28/ 30.1%	65/ 69.9%	93/100.0%
740-759	Congenital anomalies	5/ 35.8%	9/ 64.2%	14/100.0%
760-779	Certain causes of perinatal morbidity and mortality	0/ 0.0%	5/100.0%	5/100.0%
780-796	Symptoms and illdefined conditions	5/ 35.8%	9/ 64.2%	14/100.0%
800-999	Accidents, poisonings, and violence	9/ 40.9%	13/ 59.1%	22/100.0%
	Total	375/ 32.1%	793/ 67.9%	1168/100.0%
	Number of patients	127	370	

TABLE XXV

COMPARISON OF OBESITY DIAGNOSIS.

Physician	Number of Patients	Number Diagnosed as Obese	Percent Diagnosed as Obese
1	154	22	14.2
2	414	95	22.9
3	608	103	16.9
4	344	6	1.7
5	457	177	38.7
Total	1977	403	20.4

To this point no mention has been made during the test and evaluation concerning the possible diseases or abnormalities that the 28 variables could possible detect. In fact, to this point in the test and evaluation, it has been assumed that the variables could detect all diseases.

Quite obviously the model is not capable of detecting all the one thousand diagnosis codes (Reference 19). In fact only a small number of the diseases are detectable when a diagnosis is based solely on the 28 variables. After discussion with the medical staff, the following

disease classes were decided as detectable by the model's variables:

1. Diseases of thyroid gland 240-246
2. Diseases of other endocrine glands 250-258
3. Avitaminoses and other nutritional deficiency 260-269
4. Other metabolic diseases (excluding obesity) 270-279
5. Diseases of the blood and blood forming organs 280-289
6. Active rheumatic fever 390-392
7. Chronic rheumatic heart 393-398
8. Hypertensive disease 400-404
9. Ischemic heart disease 410-414
10. Other forms of heart disease 420-429
11. Diseases of veins and lymphatics, and other diseases of circulatory system 450-458
12. Pneumonia 480-486
13. Bronchitis, emphysema, and asthma 490-493
14. Other diseases of respiratory system 510-519
15. Appendicitis 540-543
16. Diseases of liver, gallbladder, and pancres 570-577
17. Nephartis and nephrosis 580-584
18. Other diseases of urinary system 590-599
19. Internal injury of chest, abdomen, and pelvis 860-869
20. Adverse effects of medicinal agents 960-979.

The sample of 923 patients is then screened to remove those patients who do not have a diagnosis corresponding to one of the classes listed above. In addition, patients who have only an obesity diagnosis and no other diagnosis are removed from the sample.

Based on this screening the sample reduced to 743 patients. The revised model's classification is given in Table XXVI. As expected, these values vary slightly from

the original sample. From the table 52.5 percent of the patients have one or more clinical value falling outside the 95 percent confidence limits. Also 18.1 percent have a chi-square value greater than the theoretical chi-square value for a level of significance of $\alpha = 0.05$ and the appropriate degrees of freedom.

TABLE XXVI

RESULTS OF MODEL'S CLASSIFICATION USING ORIGINAL SAMPLE (BASED ON ONLY THOSE DIAGNOSES POSSIBLE FROM THE 28 VARIABLES)

		95 Percent Multivariate Chi-square Test		
		Normal	Abnormal	Total
95 Percent Univariate Confidence Limits	Within	342/46.0%	11/ 1.5%	353/ 47.5%
	Outside	267/35.9%	123/16.6%	390/ 52.5%
	Total	609/81.9%	134/18.1%	743/100.0%

The revised comparison of the model's classification with the doctor's diagnosis is given in Table XXVII. From the table for 63.1 percent of the sample the model agreed with the doctor and for 36.9 percent the model disagreed with the doctor. There is no noticeable difference in these percentages with the percentages in Table XXII. A similar test for testing the dependence between the model's classification and the doctor's diagnosis indicates that there is a dependence.

TABLE XXVII

MODEL'S CLASSIFICATION COMPARED WITH DOCTOR'S DIAGNOSIS (BASED ON ONLY THOSE DIAGNOSES POSSIBLE FROM THE 28 VARIABLES

		Doctor's Diagnosis		
		Normal	Abnormal	Total
Model's Classification	Normal	247/33.2%	95/12.8%	342/ 46.0%
	Abnormal	179/24.1%	222/29.9%	401/ 54.0%
	Total	426/57.3%	317/42.7%	743/100.0%

Table XXVIII gives an indication as to the accuracy of the model in detecting particular diseases. Based on the doctor's diagnosis, the sample of 743 patients consists of 426 normal and 317 abnormal patients. The 317 abnormal patients have 376 diagnoses which could possibly be detected by the model's variables. The model successfully classified 222 of the 317 (70.2 percent) as abnormal. These 222 patients have 261 diagnosis codes or 69.5 percent of the 376 diagnoses for the sample.

For those diseases having a frequency of ten or more, the model was able to detect over seventy percent of the abnormalities for diagnosis codes 240-246, diseases of the thyroid gland; 250-258, diseases of other endocrine glands; 400-404, hypertension; and 450-458, diseases of veins and lymphatics, and other diseases of the circulatory system. The model was also able to detect over sixty percent of the abnormalities for diagnosis codes 410-414, ischemic heart disease; and 420-429, other forms of heart disease.

Of the eight diagnoses for a disease of the respiratory system (510-519), the model was only able to detect three. After discussion with the medical staff, it was concluded that the adding of several additional variables should reduce this error. The first variables expresses FEV_1 as a percentage of vital capacity. The second variable

TABLE XXVIII

COMPARISON OF MODEL'S CLASSIFICATION VERSUS PATIENT DIAGNOSIS (USING INITIAL SAMPLE)

Major Disease Group		Number of Patients		Total
		Normal by Model and Abnormal by Doctor	Abnormal by Both Model and Doctor	
240-246	Diseases of thyroid gland	3/ 30.0%	7/ 70.0%	10/100.0%
250-258	Diseases of other endocrine glands	9/ 26.4%	25/ 73.6%	34/100.0%
260-269	Avitaminoses and other nutritional deficiency	1/ 50.0%	1/ 50.0%	2/100.0%
270-279	Other metabolic diseases (excluding obesity)	46/ 38.3%	74/ 61.7%	120/100.0%
280-289	Diseases of the blood and blood forming organs	0/ 0.0%	1/100.0%	1/100.0%
390-392	Active rheumatic fever	0/ 0.0%	0/ 0.0%	0/ 0.0%
393-398	Chronic rheumatic heart	0/ 0.0%	2/100.0%	2/100.0%
400-404	Hypertensive disease	19/ 23.2%	63/ 76.8%	82/100.0%
410-414	Ischemic heart disease	16/ 34.1%	31/ 65.9%	47/100.0%
420-429	Other forms of heart disease	4/ 30.8%	9/ 69.2%	13/100.0%
450-458	Diseases of veins and lymphatics, and other diseases of circulatory system	11/ 24.5%	34/ 75.5%	45/100.0%
480-486	Pneumonia	0/ 0.0%	0/ 0.0%	0/ 0.0%
490-493	Bronchitis, emphysema, and asthma	1/ 12.5%	7/ 87.5%	8/100.0%
510-519	Other diseases of respiratory system	5/ 62.5%	3/ 37.5%	8/100.0%
540-543	Appendicitis	0/ 0.0%	0/ 0.0%	0/ 0.0%
570-577	Diseases of liver, gallbladder, and pancres	0/ 0.0%	3/100.0%	3/100.0%
580-584	Nephritis and nephrosis	0/ 0.0%	1/100.0%	1/100.0%
590-599	Other diseases of urinary system	0/ 0.0%	0/ 0.0%	0/ 0.0%
860-869	Internal injury of chest, abdomen and pelvis	0/ 0.0%	0/ 0.0%	0/ 0.0%
960-979	Adverse effects of medicinal agents	0/ 0.0%	0/ 0.0%	0/ 0.0%
Total		115/ 30.8%	261/ 69.2%	376/100.0%

expresses vital capacity as a percentage of the predicted based on the patient's height, weight, and frame.

For the remaining ten diagnosis categories the frequency of occurrence of the diseases in the sample is too small to test the model's capability of detecting those diseases.

The distribution of the number of values falling outside the 95 percent confidence limits is given in Table XXIX. The 899 values falling outside the limits represent 538 patients. Of the values falling outside the limits, 470, or 52.2 percent are greater than two sigma while 429, or 47.8 percent, are less than two sigma. The majority of the blood pressure, bilirubin, red and white cell, and urine pH values fell outside the greater than two sigma limits. Also the majority of the vital capacity and FEV_1 values fell outside the less than two sigma limits.

Analysis of Misclassification

In comparing the results of the model's classification with the doctor's diagnosis, for 12.8 percent of the sample the model classified the patients normal while the doctor diagnosed them abnormal. This percentage can be considered the actual error of the model.

TABLE XXIX

VALUES FALLING OUTSIDE LIMITS

Variable	Number Minus Two Sigma	Number Plus Two Sigma	Total	Percent of Grand Total
Actual/Ideal Weight	17	0	17	2
Pulse	12	20	32	4
Systolic Blood Press	8	27	35	4
Diastolic Blood Press	8	23	31	3
Arm Skin Folds	18	11	29	3
Back Skin Folds	19	9	28	3
Hematocrit	31	19	50	6
White Blood Count	19	10	29	3
Glucose	12	15	27	3
Cholesterol	27	9	36	4
Uric Acid	36	25	61	7
SGPT	11	3	14	2
Bilirubin, Total	0	12	12	1
Urine Red Cells	0	26	26	3
Urine White Cells	0	9	9	1
Specific Gravity	21	7	28	3
Urine pH	0	43	43	5
Total Heart Diameter	21	31	52	6
Thoracic Diameter	26	18	44	5
T.D./TH.D.	18	30	48	5
EKG Heart Rate	24	26	50	6
PR Interval	26	32	58	6
QRS Duration	13	38	51	6
QRS Axis	28	12	40	4
Vital Capacity	15	7	22	2
Forced Expiration Vol	19	8	27	3
Total	429	470	899	100

The other error between the model's classification and the doctor's diagnosis is that the model classified 24.1 percent of the sample abnormal while the doctor diagnosed them normal. This error is not such a serious error as the previous error. The results of this error are that more patients would be referred for unnecessary additional treatment.

However, on the other hand, it may be that the doctor failed to detect an abnormality in these patients. And, since the model considers the correlation pattern between the variables, it may be that the model is able to detect an abnormality which is overlooked by the physician. To check this possibility, the 179 patients, who are classified as abnormal by the model and as normal by the doctor, are divided into two groups: those failing only the univariate test and those failing both the univariate and multivariate tests. The possibility of a patient being abnormal is higher for those who failed both of the tests. Therefore, from these 48 patients, a random sample of 15 patients was selected for detailed analysis.

Of these 15 patients, ten were considered normal by the physician, two had a bad test value, and three were considered abnormal. Of the three abnormal patients, one was diagnosed as having hypertension, one was diagnosed as a

high risk because of a low FEV_1, and one had an abnormal electrocardiogram which probably resulted from a mild heart attack.

Using this sample of 15 patients, three patients, or 20 percent are now diagnosed as abnormal rather than normal. Therefore, the model's classification of 24.1 percent of the patients as abnormal while the doctor diagnosed them normal is reduced to 19.4 percent. Likewise, the model and the doctor both classified 34.6 percent of the patients as abnormal as compared to 29.9 percent. This increases the agreement between the model and the doctor to 67.8 percent.

Analysis Using Independent Sample

The second step in the analysis consists of taking a random sample from an independent data source. Since the original sample consisted of physicals given through September, 1971, an independent sample would be those exams given after October 1, 1971. This sample consists of over 200 patients. After eliminating those physicals having diagnoses which are not detectable by the model's variables, the sample reduces to 174 patients.

Model's Classification

The results of the model's classification are given in Table XXX. From the table 65.5 percent of the patients have one or more clinical variable falling outside the 95 percent confidence limits. On the other hand, 22.3 percent of the patients have a chi-square value greater than the theoretical chi-square for $\alpha = 0.05$ and the appropriate degrees of freedom. Also 1.1 percent passed the univariate test but failed the multivariate test, while 44.3 percent failed the univariate test but passed the multivariate test.

TABLE XXX

RESULTS OF MODEL'S CLASSIFICATION USING INDEPENDENT SAMPLE

		95 Percent Multivariate Chi-square Test		
		Normal	Abnormal	Total
95 Percent Univariate Confidence Limits	Within	58/33.4%	2/ 1.1%	60/ 34.5%
	Outside	77/44.3%	37/21.2%	114/ 65.5%
	Total	135/77.7%	39/22.3%	174/100.0%

Comparing the independent sample with the initial sample, it is seen that the independent sample classified fewer patients as passing both the univariate and multivariate tests (33.4 versus 46.0). Likewise, more patients failed both the two tests (21.2 versus 16.6).

The error rates are similar between the two samples. For the independent sample the model classified 44.3 percent as failing the univariate but passing the multivariate tests as compared to 35.9 percent for the initial sample. Also, for the independent sample, the model classified 1.1 percent as passing the univariate but failing the multivariate test as compared to 1.5 percent for the initial sample.

The cumulative distribution of the chi-square values for those patients having values for all twenty-eight variables is given in Figure 19. The dotted curve is the cumulative distribution of the initial sample. Since there are more abnormal patients in the independent sample, the chi-square cumulative distributions do not exactly agree.

Model's Classification Versus Doctor's Diagnosis

The comparison of the model's classification with the doctor's diagnosis is given in Table XXXI. The larger percentages of abnormal diagnoses by the doctor confirms the results of the previous paragraph; that the independent

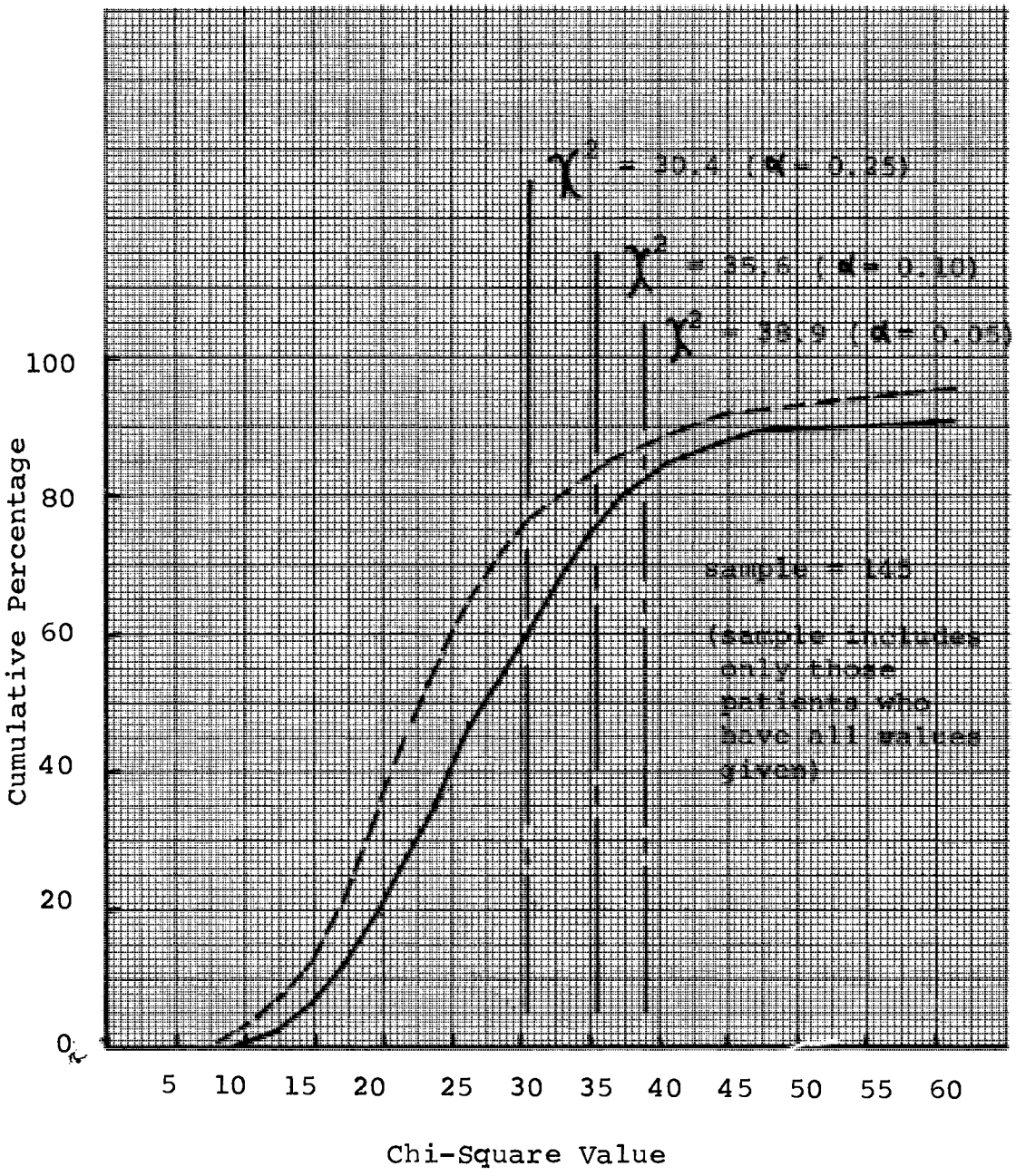


Figure 19. Distribution of Chi-Square Values for Independent Sample

sample has a relative larger number of abnormal patients as compared to the initial sample.

TABLE XXXI

MODEL'S CLASSIFICATION COMPARED WITH DOCTOR'S DIAGNOSIS (USING INDEPENDENT SAMPLE)

		Doctor's Diagnosis		
		Normal	Abnormal	Total
Model's Classification	Normal	35/20.1%	23/13.2%	58/ 33.3%
	Abnormal	44/25.3%	72/41.4%	116/ 66.7%
	Total	79/45.4%	95/54.6%	174/100.0%

From the table 20.1 percent of the patients who are diagnosed as normal by the doctor are also diagnosed as normal by the model. Likewise, 41.4 percent of the patients who are diagnosed as abnormal by the doctor are also classified as abnormal by the model. This is equivalent of saying that for 61.5 percent of the sample the model agreed with the doctor and for 38.5 percent the model disagreed with the doctor. This compares favorably with the initial sample of 66.9 percent and 33.1 percent, respectively.

The test for agreement between the model's results and the doctor's diagnosis gives a calculated chi-square value of 7.72 which is greater than the theoretical chi-square value of 3.84 for $\alpha = 0.05$ and one degree of freedom. Therefore, there is a dependence between the model's classification and the doctor's diagnosis.

The error rates are similar between the two samples. For the independent sample the model classified 25.3 percent of the patients abnormal while the doctor diagnosed them normal. This compares with 24.1 percent for the initial sample. This increase is not critical since the only error is that possibly the patients may receive additional medical attention which is not required. In addition, the model classified 13.2 percent as normal while the doctor diagnosed them abnormal. This compares with 12.8 percent for the initial sample.

Model's Classification Versus Patient Diagnosis

Table XXXII gives an indication as to the accuracy of the model in detecting specific diseases. Table XXXIII gives a comparison between the independent and initial samples. Although the sample size is smaller for the independent sample, there is close agreement between the two. In fact, based on percentages, the model was able to detect

TABLE XXXII

COMPARISON OF MODEL'S CLASSIFICATION VERSUS PATIENT DIAGNOSIS (USING INDEPENDENT SAMPLE)

Major Disease Group		Number of Patients		Total
		Normal by Model and Abnormal by Doctor	Abnormal by Both Model and Doctor	
240-246	Diseases of thyroid gland	2/ 66.7%	1/ 33.3%	3/100.0%
250-258	Diseases of other endocrine glands	0/ 0.0%	9/100.0%	9/100.0%
260-269	Avitaminoses and other nutritional deficiency	0/ 0.0%	1/100.0%	1/100.0%
270-279	Other metabolic diseases (excluding obesity)	15/ 26.3%	42/ 73.7%	57/100.0%
280-289	Diseases of the blood and blood forming organs	0/ 0.0%	0/ 0.0%	0/ 0.0%
390-392	Active rheumatic fever	0/ 0.0%	0/ 0.0%	0/ 0.0%
393-398	Chronic rheumatic heart	0/ 0.0%	2/100.0%	2/100.0%
400-404	Hypertensive disease	1/ 3.6%	27/ 96.4%	28/100.0%
410-414	Ischemic heart disease	4/ 36.4%	7/ 63.6%	11/100.0%
420-429	Other forms of heart disease	2/ 33.3%	4/ 66.7%	6/100.0%
450-458	Diseases of veins and lymphatics, and other diseases of circulatory system	1/ 16.7%	5/ 83.3%	6/100.0%
480-486	Pneumonia	0/ 0.0%	0/ 0.0%	0/ 0.0%
490-493	Bronchitis, emphysema, and asthma	1/ 25.0%	3/ 75.0%	4/100.0%
510-519	Other diseases of respiratory system	0/ 0.0%	1/100.0%	1/100.0%
540-543	Appendicitis	0/ 0.0%	1/100.0%	1/100.0%
570-577	Diseases of liver, gallbladder, and pancres	0/ 0.0%	0/ 0.0%	0/ 0.0%
580-584	Nephritis and nephrosis	0/ 0.0%	0/ 0.0%	0/ 0.0%
590-599	Other diseases of urinary system	0/ 0.0%	0/ 0.0%	0/ 0.0%
860-869	Internal injury of chest, abdomen, and pelvis	0/ 0.0%	0/ 0.0%	0/ 0.0%
960-979	Adverse effects of medicinal agents	0/ 0.0%	0/ 0.0%	0/ 0.0%
	Total	26/ 20.2%	103/ 79.8%	129/100.0%

TABLE XXXIII

COMPARISON BETWEEN INITIAL AND INDEPENDENT SAMPLES

	Diagnosis Group	Initial Sample			Independent Sample		
		Size	Correct Diagnoses	Percent Correct	Size	Correct Diagnoses	Percent Correct
240-246	Diseases of thyroid gland	10	7	70.0	3	1	33.3
250-258	Diseases of other endocrine glands	34	25	73.6	9	9	100.0
260-269	Avitaminoses and other nutritional deficiency	2	1	50.0	1	1	100.0
270-279	Other metabolic diseases excluding obesity	120	74	61.7	57	42	73.7
280-289	Diseases of the blood and blood forming organs	1	1	100.0	0	0	
390-392	Active rheumatic fever	0	0		0	0	
393-398	Chronic rheumatic heart	2	2	100.0	2	2	100.0
400-404	Hypertensive disease	82	63	76.8	28	27	96.4
410-414	Ischemic heart disease	47	31	65.9	11	7	63.6
420-429	Other forms of heart disease	12	9	66.7	6	4	66.7
450-458	Diseases of veins and lymphatics, and other diseases of circulatory system	45	34	75.7	6	5	83.3
480-486	Pneumonia	0	0		0	0	
490-493	Bronchitis, emphysema, and asthma	8	7	87.5	4	3	75.0
510-519	Other diseases of respiratory system	8	3	37.5	1	1	100.0
540-543	Appendicitis	0	0		1	1	100.0
570-577	Diseases of liver, gallbladder, and pancres	3	3	100.0	0	0	
580-584	Nephritis and nephrosis	1	1	100.0	0	0	
590-599	Other diseases of urinary system system	0	0		0	0	
860-869	Internal injury of chest, abdomen, and pelvis	0	0		0	0	
960-979	Adverse effects of medicinal agents	0	0		0	0	
	Total	376	261	69.2	129	103	79.8

more of the diagnoses in the independent sample. This is especially true for other metabolic diseases and hypertension. From the independent sample, 73.7 percent of the other metabolic diseases were detected as compared with 61.7 percent for the initial sample. Also, 96.4 percent of the hypertension diagnoses were detected for the independent sample as compared with 76.8 percent for the initial sample.

Ranking of Major Contributors

As part of the patient profile, it is of interest to identify those variables which are the major contributors to the patient's chi-square value. Then, once these variables are identified, are these the variables that correspond to the doctor's diagnosis? For example, a patient diagnosed as having hypertension should have a high blood pressure. Then, did the model identify blood pressure as a major contributor to the patient's chi-square?

To check the above question, a sample was collected from those patients who are diagnosed as abnormal by both the model and the doctor. A count is then made of the frequency that each variable is one of the three major contributors. These results are presented in Table XXXIV.

TABLE. XXXIV

MAJOR VARIABLES CONTRIBUTING TO CHI-SQUARE VALUE VERSUS DOCTOR DIAGNOSIS

Variable	Diagnosis Group 240-246	250-258	260-269	270-279	280-289	390-392	393-398	400-404	410-414	420-429	450-458	480-486	490-493	510-519	540-543	570-577	580-584	590-599	860-869	960-970	Total
Actual/Ideal Weight																					0
Pulse	1	4		3				6	2		3		1								20
Systolic Blood Press	1	9		8	1			21	2	2	5			1							50
Diastolic Blood Press	1	2		9			1	20	6	1	6										46
Arm Skin Folds		1		2			1	9	1		5		1								20
Back Skin Folds	1	1	1	9				4		3	2										21
Hematocrit	3	3		3	1			7	2	1	3										23
White Blood Count		3		12			1	6			5		1								28
Glucose		9		4			1	7	3		2										26
Cholesterol		1		15				6	4	3	4					1					34
Uric Acid		5		20				14	4		7			1							51
SGPT	1			3				3	3												10
Bilirubin, Total		2		7				2	1		2		1			1					16
Urine Red Cells				4				4	3		2			1				1			15
Urine White Cells				3					3												6
Specific Gravity		5		6				4	1		1										17
Urine pH	1	3		12			1	8	3	1	6			2		1					38
Total Heart Diameter	3	5		14	1			10	7	2	4		1	1		1					49
Thoracic Diameter	1	2	1	16				9	5	1	10		3	1							49
T.D./TH.D.	2	8	1	14				12	11	4	10		5					1			68
EKG Heart Rate	1	5		12				12	6	2	3		1			3					45
PR Interval		3		8				3	4	1	2		1								22
QRS Duration	3	2		7				8	3	3	1		2			2					31
QRS Axis	2	2		6			1	7	4	2	1		1	1				1			28
Vital Capacity				2				5	2	1	3		1	1							15
Forced Expiration Vol				2				2	4		3		2								13
Total	21	75	3	201	3	0	6	189	84	27	90	0	21	9	0	9	0	3	0	0	741
Number of Patients	7	25	1	67	1	0	2	63	28	9	31	0	7	3	0	3	0	1	0	0	247

For example, of the 25 patients diagnosed as having diabetes, nine of the patients had profiles which indicated glucose as a major contributor. A review of the table with the medical staff resulted in one comment: "interesting."

Analysis of Longitudinal Drift

To adequately analyze the chi-square values as a composite measure of the patient's health, more than one physical examination is needed per patient. However, since the Medical Center modified the data elements in January, 1971, only a small percentage of the patients presently have more than one exam which contains measurements for the selected twenty-eight variables. In fact, only 11 patients have more than one of these exams.

A plot of the longitudinal drift for this sample of 11 patients is given in Figure 20. Only two of the patients have an increased chi-square value. Three patients who were classified as abnormal during their first examination are now classified normal (using $\alpha = 0.05$). One of these three patients had a very high urine white cell count on his first exam and a normal count on his second. Another of the three had a high systolic blood pressure on his first exam and a more normal value on his second. The third patient's hematocrit was low on his first exam and normal on his second.

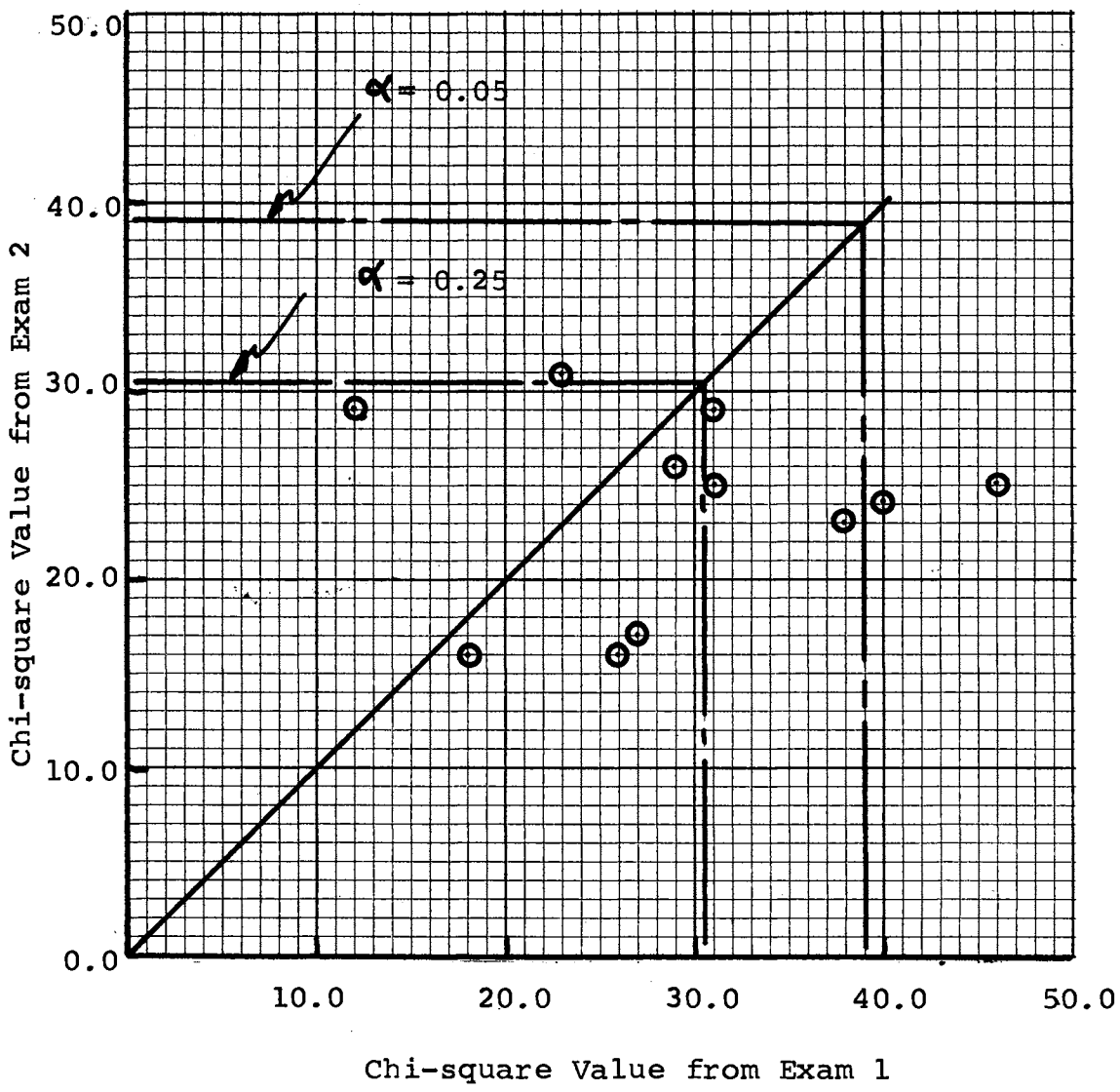


Figure 20. Longitudinal Drift in Patient Chi-Square Values

Because of the small sample no conclusion could be drawn at this time concerning the chi-square value as a possible measure of the drift in a patient's health over time.

Sources of Error

During the test and evaluation several sources of error were detected. These errors are briefly discussed in the following paragraphs.

One source of error resulted from classifying patients who are presently taking prescribed medications to control abnormalities. It was revealed that the doctors are diagnosing these patients as abnormal even though they are taking medication to control the abnormality. This error is most noticeable for diabetes and hypertension. By taking the proper medication the patient can control his glucose and blood pressure.

Another source of error is the exclusion of several variables which are a function of vital capacity (VC) and FEV_1. The first variable expresses FEV_1 as a percentage of total capacity: $FEV_1/VC \times 100$. The second variable expresses vital capacity as a percentage of the predicted based on the patient's height, weight, and frame. The

adding of these two variables should increase the model's capability of detecting respiratory abnormalities.

Another source of error is that not all the physicians use the same criteria for diagnosing abnormalities. This error is most noticeable in diagnosing obesity; therefore, as a result, this diagnosis was removed from the analysis. A check of the other diagnoses did not reveal any other major discrepancies between the physicians. However, the diagnosing of hypertension did vary between doctors. The prevalence of hypertension for the total population is 9.0 percent. The prevalence among the doctor's diagnoses are 15.6, 12.3, 4.6, 9.0, and 9.9.

Another source of error is that the model weighs low values as significantly as high values. For example, in most instances, a low cholesterol is much better than a high cholesterol. However, in computing the chi-square value, a low cholesterol is weighed the same as a high cholesterol. Conversely, higher vital capacities and FEV_1's are more desirable than lower values. This type of error is quite common and is generally ignored by assuming that the errors will tend to cancel each other. By adding the two additional pulmonary variables some of this error would be removed.

A problem encountered in the testing of the model is that, after adjusting the variables conditionally for the patient's height and weight, the variable actual/ideal weight had almost no significance. This is apparent by observing that no actual/ideal weight values fell outside the 95 percent tolerance region. One approach to solving this problem is to make the variables conditional only to height and then remove weight from the list of variables.

CHAPTER V

CONCLUSIONS

Summary of Results

The results of this research is a tool to assist the physician in analyzing a patient's health. Emphasis is placed on the periodic physical examination as the principle source of data input. For this reason the results of this research are more applicable to health clinics and industrial medical centers where patients go for periodic physicals or for some type of screening, than to private physicians who are generally concerned with treating existing ailments.

To assist the physician in analyzing a patient's health, three outputs have been developed. Computer programs have been written for generating these three outputs.

The principle output is the patient health profile. This output displays twenty-six of the patient's clinical variables with respect to established norms for the patient age group. Those variables falling outside the 95 percent confidence limits are flagged for further medical attention.

A multivariate chi-square test is also part of the health profile. The computed chi-square value, which considers the correlation pattern between the variables, can be considered as a composite measure of the patient's health.

The second output is a plot of the composite measure of the patient's health (i.e., his chi-square values) as a function of time; that is, from exam to exam. By observing this composite measure, it is possible to detect any overall longitudinal drift in the patient's health. Various levels of significance can be assigned to these composite measures for assigning risk factors to a possible drift in the patient's health.

The third output is a summary of the patient's medical folder. Included in this summary are a history of the patient's prior physical and laboratory details, and all his prior scheduled and nonscheduled visits for some type of medical treatment.

The model's capability of detecting abnormalities was tested by taking a sample of patients from the initial sample and from an independent source. The results are summarized in Table XXXV. From the table, using the data from the initial sample, 33.2 percent of the patients who are diagnosed as normal by the doctor are also classified as normal by the model. Likewise, 29.9 percent who are

TABLE XXXV

COMPARISON OF MODEL'S AGREEMENT

Initial Sample of 743

		Doctor's Diagnosis		
		Normal	Abnormal	Total
Model's Classification	Normal	33.2%	12.8%	46.0%
	Abnormal	24.1%	29.9%	54.0%
	Total	57.3%	42.7%	100.0%

Independent Sample of 174

		Doctor's Diagnosis		
		Normal	Abnormal	Total
Model's Classification	Normal	20.1%	13.2%	33.3%
	Abnormal	25.3%	41.4%	66.7%
	Total	45.4%	54.6%	100.0%

diagnosed as abnormal by the doctor are also classified as abnormal by the model.

The model classified 12.8 percent of the patients normal while the doctor diagnosed them abnormal. In addition, the model classified 24.1 percent of the patients abnormal while the doctor diagnosed them normal. It is this area, those patients diagnosed as normal by the doctor and abnormal by the model, where the real payoff of the model exists. It may be that for a percentage of these patients, the doctor failed to detect an abnormality. Since the model considers the correlation pattern between the variables, it may be that the model is able to detect an abnormality which was overlooked by the physician. Further analysis of the 24.1 percent resulted in 20 percent of these patients being diagnosed as abnormal.

Use of Research Results

As originally stated, the source of data for this research is the MSFC Medical Center. Therefore, many of the variables used in the development of the patient's health profile may be unique to this Center. In addition, the three outputs may also reflect some of the unique requirements and features of the Medical Center.

To date the Medical Center has begun to implement the results of this research. First, the means and standard deviations for the age groups are being used by the Medical Center as indicators of the distributions of the variables of the MSFC population.

The second use of the results of this research is that the patient summaries are being added to the patient folders. The summaries are to be updated each time a patient visits the Medical Center for a physical. To date, patient summaries have been added to all employees having executive physicals.

The third use of the results of this research is the Medical Center's acceptance of the patient health profile as an excellent tool to assist the physician in analyzing a patient's health. Plans are under way to start computing health profiles on all physicals given in 1971.

Possible Uses of Research

Considerable more interest is beginning to be focused on the health of the nation's population. This is evident by the large appropriations being made to medical research by the government as well as by industries and universities.

One such area of interest which is presently being pursued by many research organizations is the development

of a mobile health clinic. Such a clinic would travel to the patient, rather than the patient traveling to the clinic. At the clinic the patient receives a routine physical examination generally administered by technicians. The clinic would be connected via telephone lines to a computer complex. At the computer complex a complete medical file would be maintained on the patient. All the analysis on the patient's physical and lab would be done by the computer.

The results of the analysis would be returned to the mobile clinic before the patient leaves. Then, should the results indicate some abnormality or near abnormal condition, the patient could be referred to a private physician for detailed individual consultation and treatment.

It is in such an environment as just described that the research in this thesis has a great potential. At the time the patient's physical and lab data is being analyzed by the computer, the data could also be impacted against predefined age group norms and a statistical analysis made.

A pilot study has been conducted concerning the feasibility of having the results of this research on-line via a terminal. The statistical routines which create the patient health profiles have been programmed to operate in an on-line environment using the Marshall Information

Retrieval and Display System (MIRADS) (20). The use of the system is depicted in Figure 21.

Input to the system is via a teletype and consists of the values for the patient's twenty-eight variables. These values are then compared against the age group norms and the results returned via the teletype. The output is the patient health profile as shown in Figures 13, 14, and 15.

Areas of Additional Research

The findings of this research readily point to many areas of further investigation. Several of these areas are briefly presented in the following paragraphs.

The first area of additional research is the adding of the two pulmonary variables and then re-evaluating the model's capability of detecting abnormalities.

An additional source of data for increasing the number of variables is from the Medical Center's ballistocardiogram and treadmill tests. Data on these tests is just beginning to be stored in the medical data base.

Because of the limited number of patients having more than one examination, the concept of using the chi-square value as a composite measure of the patient's health could not be adequately tested. A further area of research would be to obtain additional examinations on patients and

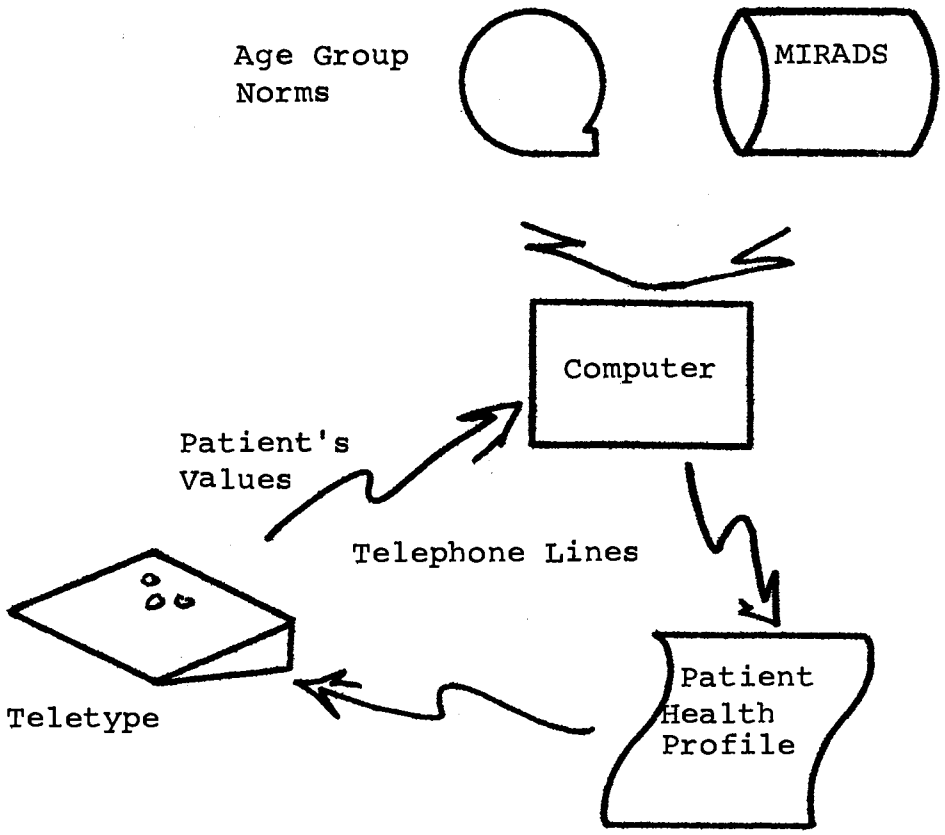


Figure 21. On-Line Capability of Patient Health Profiles

observe their longitudinal drift. It may be possible that this chi-square value could be used as an early indicator of some abnormal, or near abnormal, condition.

SELECTED BIBLIOGRAPHY

(1) Stacy, Ralph W., and Bruce D. Waxman, Computers in Biomedical Research, Vol I, Academic Press, New York, 1965.

(2) Klein, Stanley, "Automating the M.D.," Machine Design, March 19, 1970, pp 32-39.

(3) Craig, James L., and Robert L. Craig, "Mobile Multiphasic Health Testing in TVA's Occupational Health Program," Industrial Medicine, October 1970, pp 10-17.

(4) Warner, H. R., A. F. Toronto, L. G. Veasey, and R. Stephenson, "Applications to Congenital Heart Disease," Journal of American Medical Association, Vol 177, 1961.

(5) Lodwick, G. S., C. L. Haun, W. E. Smith, R. F. Keller, and E. D. Robertson, "Computer Diagnosis of Primary Bone Tumors," Radiology, Vol 80, 1963, p 273.

(6) Overall, J. E., and D. R. Gorham, "A Pattern Probability Model for the Classification of Psychiatric Patients," Behavioral Science, Vol 8, 1963, p 108.

(7) Overall, J. E., and C. M. Williams, "Conditional Probability Program for Diagnosis of Thyroid Function," Journal of American Medical Association, Vol 183, 1963, p 307.

(8) Wartak, J., "A Practical Approach to Automated Diagnosis," IEEE Transactions on Bio-Medical Engineering, Vol BME-17, No 1, January 1970.

(9) Overall, J. E., "A Configural Analysis of Psychiatric Diagnostic Stereotypes," Behavioral Science, Vol 8, 1963, p 211.

(10) Overall, J. E., and C. M. Williams, San Diego Symposium on Biomedical Engineering, Vol 13, 1963, p 141.

(11) Graybill, Franklin A., An Introduction to Linear Statistical Models, McGraw Hill, New York 1961.

(12) Gumbel, E. J., Statistics of Extremes, Columbia University Press, New York, 1958.

(13) Von Alven, William H., Reliability Engineering, Prentice Hall, Englewood Cliffs, 1964.

(14) Wilk, M. B., R. Gnanadesikan, and M. J. Huyett, "Probability Plots for the Gamma Distribution," Technometrics, Vol 4, No 1, February 1962.

(15) Edwards, A. W. F., and L. L. Cavalli-Sforza, "A Method for Cluster Analysis," Biometrics, June 1965.

(16) Morrison, D. F., Multivariate Statistical Methods, McGraw Hill, New York, 1967.

(17) Farquhar, R. D., Systemed, Inc, (personal conversation, 1970.

(18) Snedecor, George W., and William G. Cochran, Statistical Methods, The Iowa State University Press, Ames Iowa, 1967.

(19) U.S. Department of Health, Education, and Welfare, Eight Revision International Classification of Diseases, Public Health Service Publication No 1693, U.S. Government Printing Office, Washington.

(20) George C. Marshall Space Flight Center, MIRADS User's Manual, Huntsville, Alabama, 1972 (presently in printing).

APPENDIXES

APPENDIX A

DESCRIPTION OF CLINICAL VARIABLES

Physical Variables

In addition to height and weight, six physical variables are used in this research. Pulse is the rythmic beating of the arteries due to the passage of the blood waves resulting from successive contractions of the heart. The pulse is felt in the wrist by pressing the fingers on the artery.

Blood pressure is a measure of the pressure exerted by the blood on the walls of the arteries. Blood pressure is expressed in two figures: the large one, systolic, is the reading obtained at the moment when the heart contracts; and the smaller one, diastolic, is the reading at the time the heart relaxes.

The ratio actual weight/ideal weight compares the patient's actual weight with his ideal weight. The ideal weight is obtained from tables and is based on the patient's height and frame.

The arm and back skin folds are measured by pinching the back of the arm and waist and then using calipers.

Variables Related to Blood

Of the many test performed on blood, six are considered in this study. Bilirubin is derived from the hemoglobin in red blood cells which have been broken down. It is constantly being produced, and is excreted by the liver into the bile. There is always a small amount in the serum. The procedure for collecting a speciman is to withdraw venous blood and place it in a test tube and allow it to coagulate. The test is performed on the serum. To the sample of serum is added a reagent. A colored product is formed and the intensity of the color is used as a measure of the bilirubin concentration.

White blood cells are important in the defense of the body against invading microorganisms, since they destroy most harmful bacteria. The procedure for collecting a speciman is to withdraw venous blood and place the blood in a special pipette. Diluting solution is then added and the contents thoroughly mixed. The diluted suspension is then allowed to flow into a space in a special counting chamber. Through the use of a microscope, the cells per unit area are then counted and the number of cells calculated.

Cholesterol is a normal constituent of the blood and is found in all cells. In various disease states the

cholesterol in the serum may be raised or lowered. The procedure for collecting a speciman is to withdraw venous blood, place it in a test tube, and allow it to coagulate. The test is performed on the serum. Reagents are added and the color intensity is measured which is proportional to the cholesterol concentration.

The glucose test is performed to discover whether there is a disorder of glucose metabolism. The test consists of drawing venous blood. After coagulation the amount of glucose present in the serum is determined colorimetrically.

The hematocrit test measures the percent of the total volume of blood which is composed of the blood cells. The test consists of placing venous blood in an oxalate tube. The tube is then spun in a centrifuge and the height of the column of packed red blood cells measured against the graduation on the side of the tube.

The enzyme, serum glutamic pyruvic transamenase (SGPT) is found in several tissues. Its serum levels become elevated when those tissues are diseased. Amounts of SGPT present in the serum are determined colorimetrically.

The uric acid test is used to determine the uric acid concentration in the blood. The test consists of adding a reagent to a sample of blood serum which produces a blue

color with the uric acid. The intensity of the color is measured and the concentration calculated.

Variables Related to Urine

Four clinical variables relating to urine are used in this study. The Addis test consists of counting the number of cells and casts in the urine sediment. A comparison of the amount of each suggest the type of kidney disorder. The test consists of thoroughly mixing the urine speciman. A sample is then centrifuged and the sediment examined microscopically. The white cells, the red cells, and the casts are then counted.

The pH test indicates the degree of acidity of the urine. The kidney maintains the blood at the correct pH by excreting into the urine any excess ions which might alter the pH of the blood. The test consists of dropping a strip of nitrazine paper into the urine. The color change, compared to a standard chart, indicates the pH.

The specific gravity test indicates the degree of concentration of dissolved material in the urine. The test consists of placing a standard urinometer into the urine. The extent to which the urinometer sinks in the urine determines the specific gravity.

Chest X-Ray Variables

The three variables from the chest x-ray are the total heart diameter, the thoracic diameter, and the T.D./TH.D. ratio (See Figure 22). The total heart diameter and the thoracic diameter are measured directly from the x-ray using a scale.

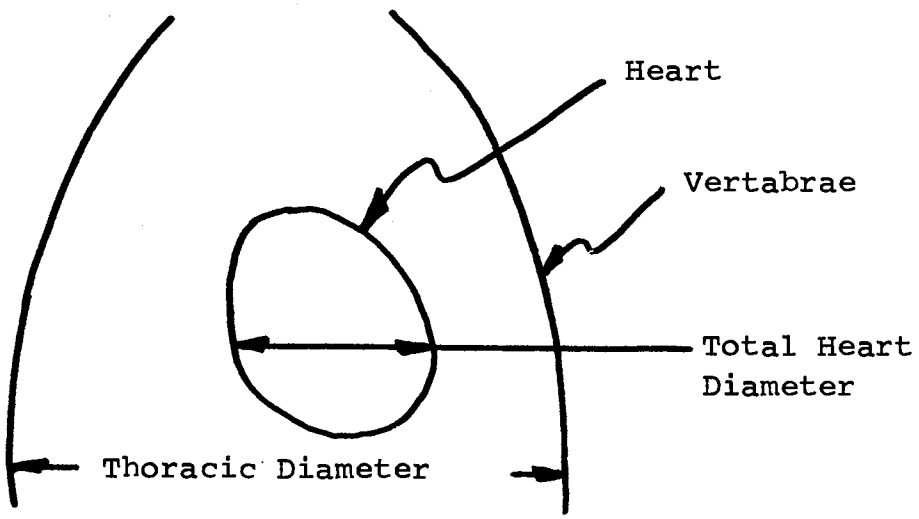


Figure 22. Chest X-Ray

Electrocardiogram Variables

The electrocardiograph is an instrument for recording the changes in the electrical potential of the heart which are transmitted through the limbs and chest wall. The electrocardiogram records the electrical potentials of the heart.

If the patient is considered as the conductor and the electrical impulses originating in the heart as the source of potential differences, then the magnitude and direction of the current produced may be measured. The typical EKG of a cardiac cycle appears as shown in Figure 23 and consists of a series of waves designated by the P wave, the QRS complex, the T wave, and the U wave.

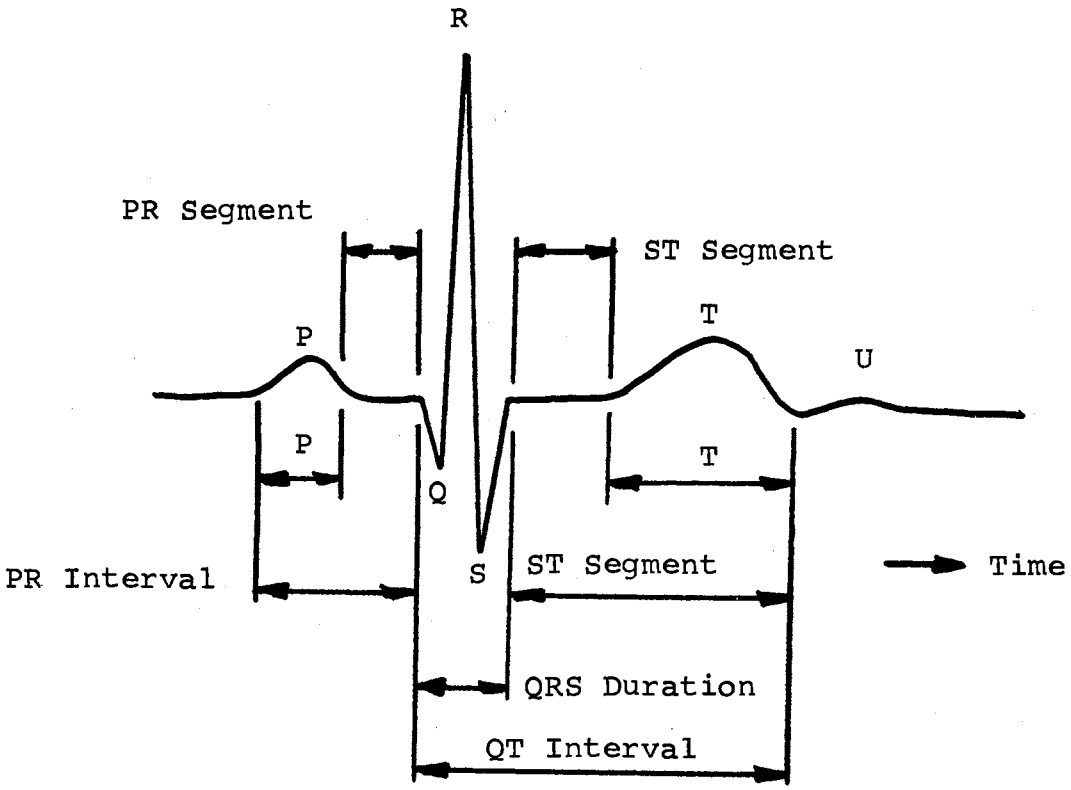


Figure 23. Waves of the Electrocardiogram

Those variables which are considered for this study are: PR interval which is measured from the beginning of the P wave to the beginning of the QRS complex; the QRS duration which is measured from the first wave of the complex to the end of the last wave of the complex; the QRS axis which is the magnitude of the mean electrical axis of the QRS complex; and the heart rate which is determined from the cycle length.

Pulmonary Function Variables

The pulmonary function tests are usefull in detecting general airway obstructions. These tests are conducted using a recording spirometer. Two variables are available from the spirometer: vital capacity and forced expiration volume.

The vital capacity test consists of the patient taking a deep breath and then blowing out slowly, and as completely as possible, all the air in the patient's lungs. The total volume delivered is called the vital capacity.

The forced expiration volume test consists of the patient taking a deep breath and then blowing out as large a volume of air as possible in one second. The total volume of air delivered in one second is called the forced expiration volume (FEV_1).

APPENDIX B

SAMPLE PLOTS OF VARIABLES ON PROBABILITY PAPER

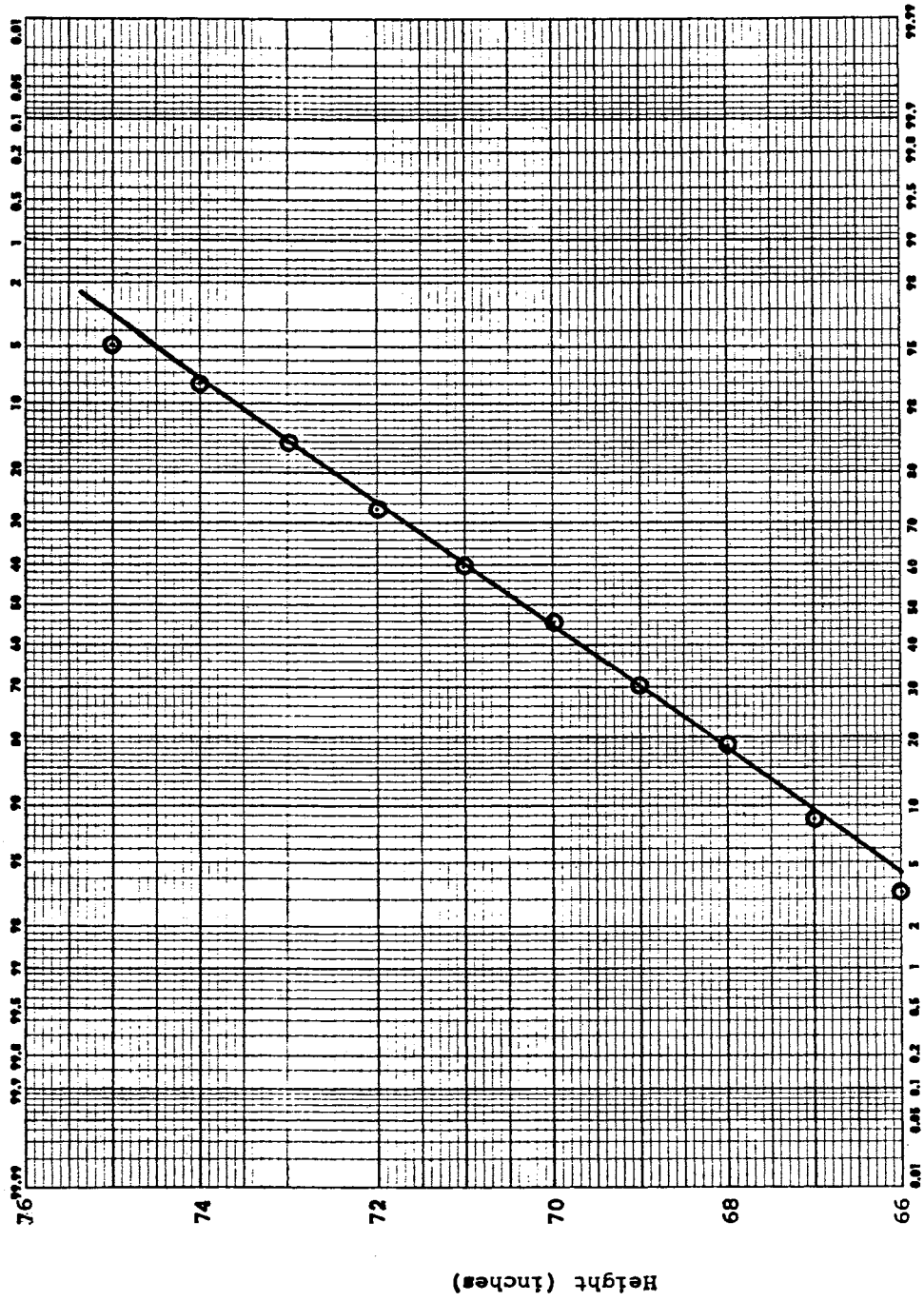


Figure 24. Height for Age Group 55-59

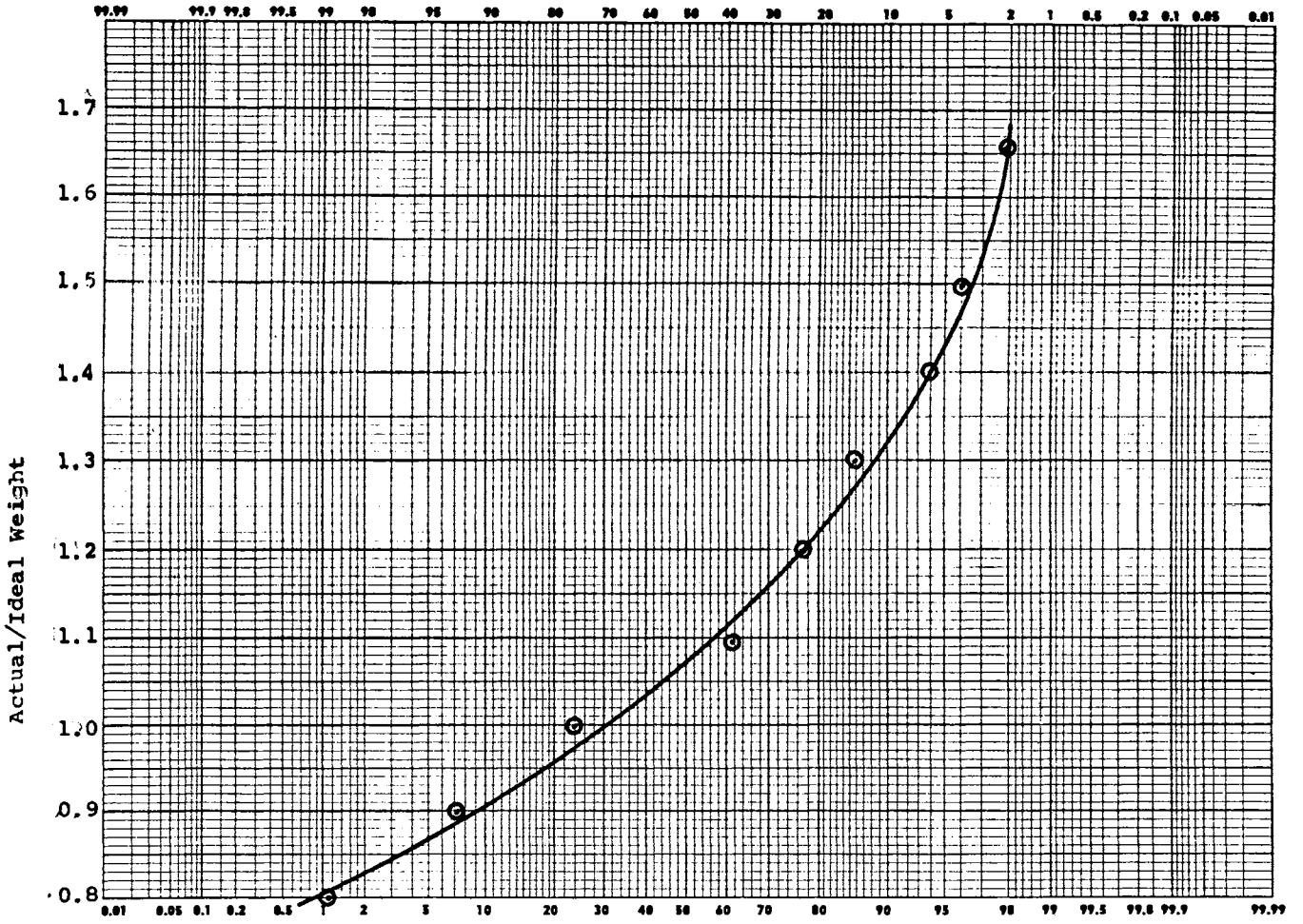


Figure 25. Actual/Ideal Weight for Age Group 55-59

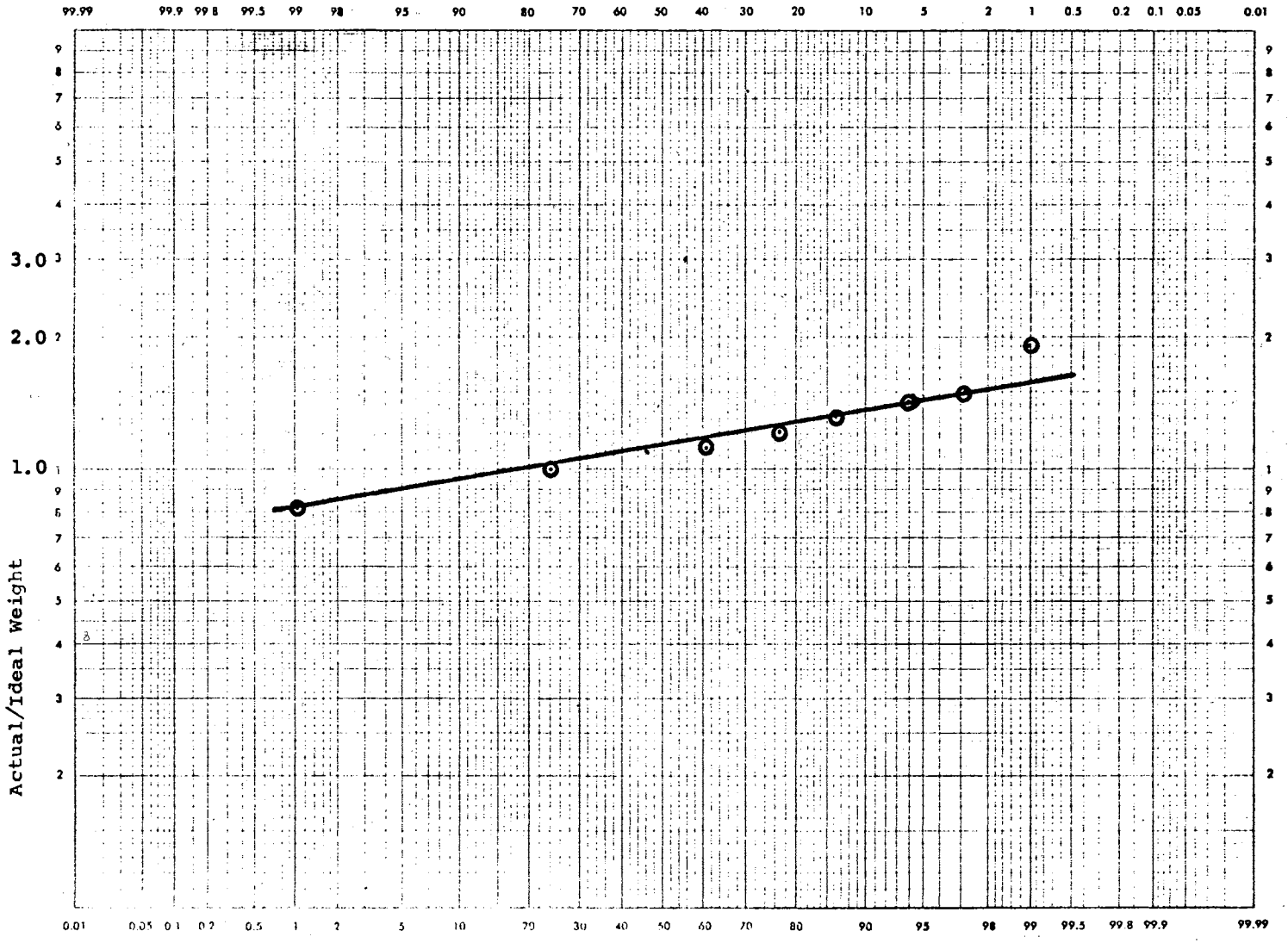


Figure 26. Actual/Ideal Weight for Age Group 55-59

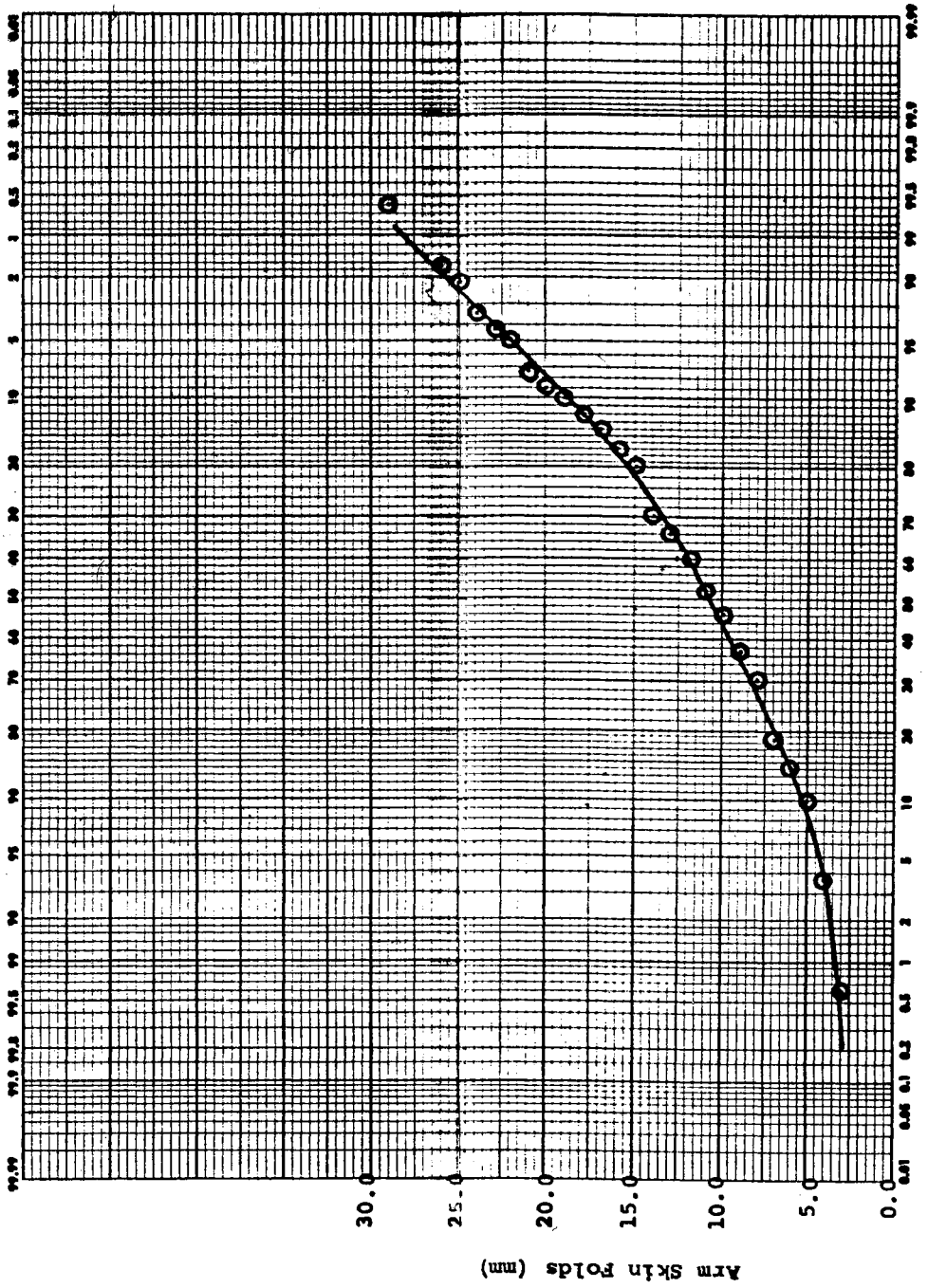


Figure 27. Arm Skin Folds for Age Group 55-59

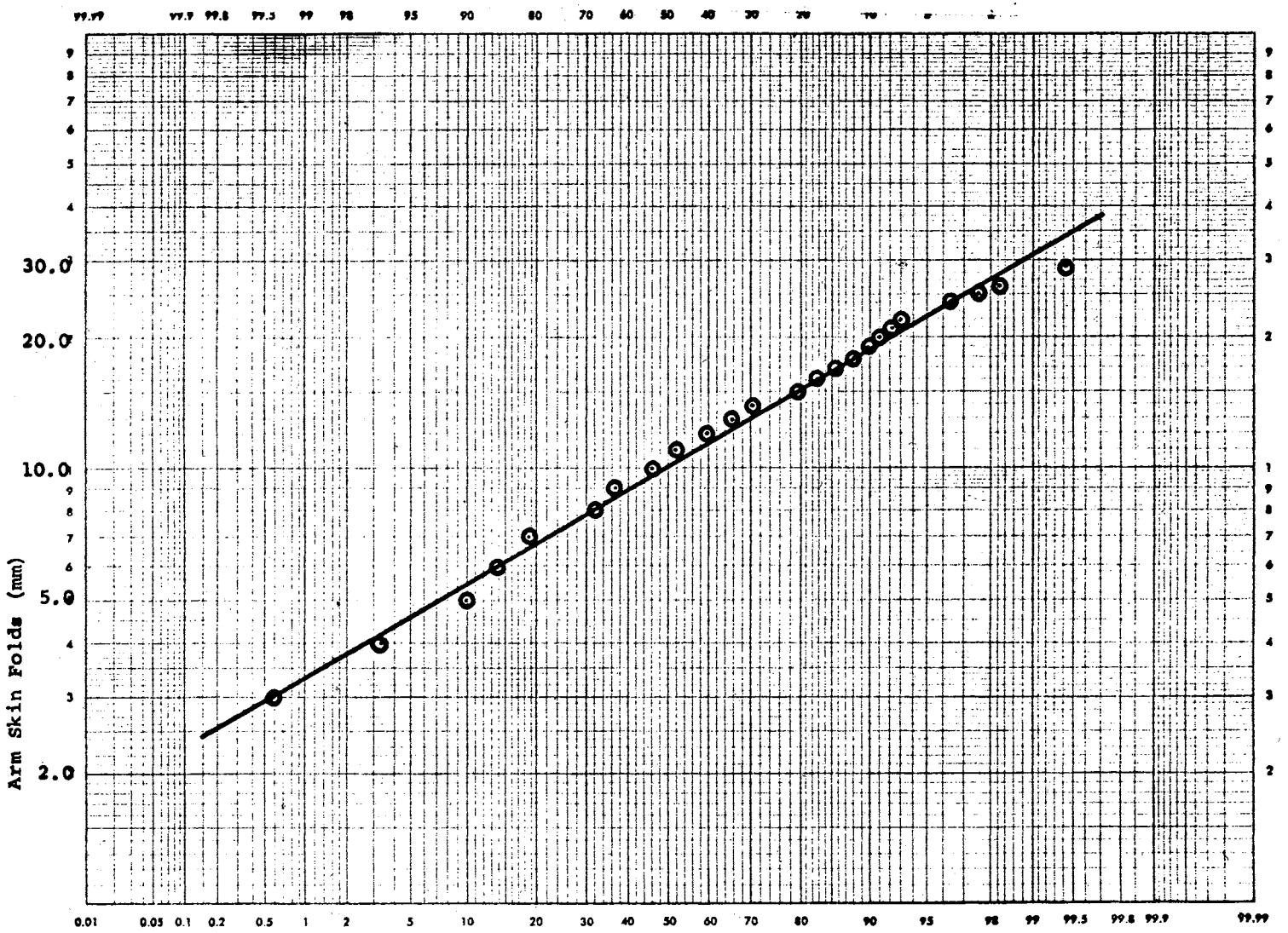


Figure 28. Arm Skin Folds for Age Group 55-59

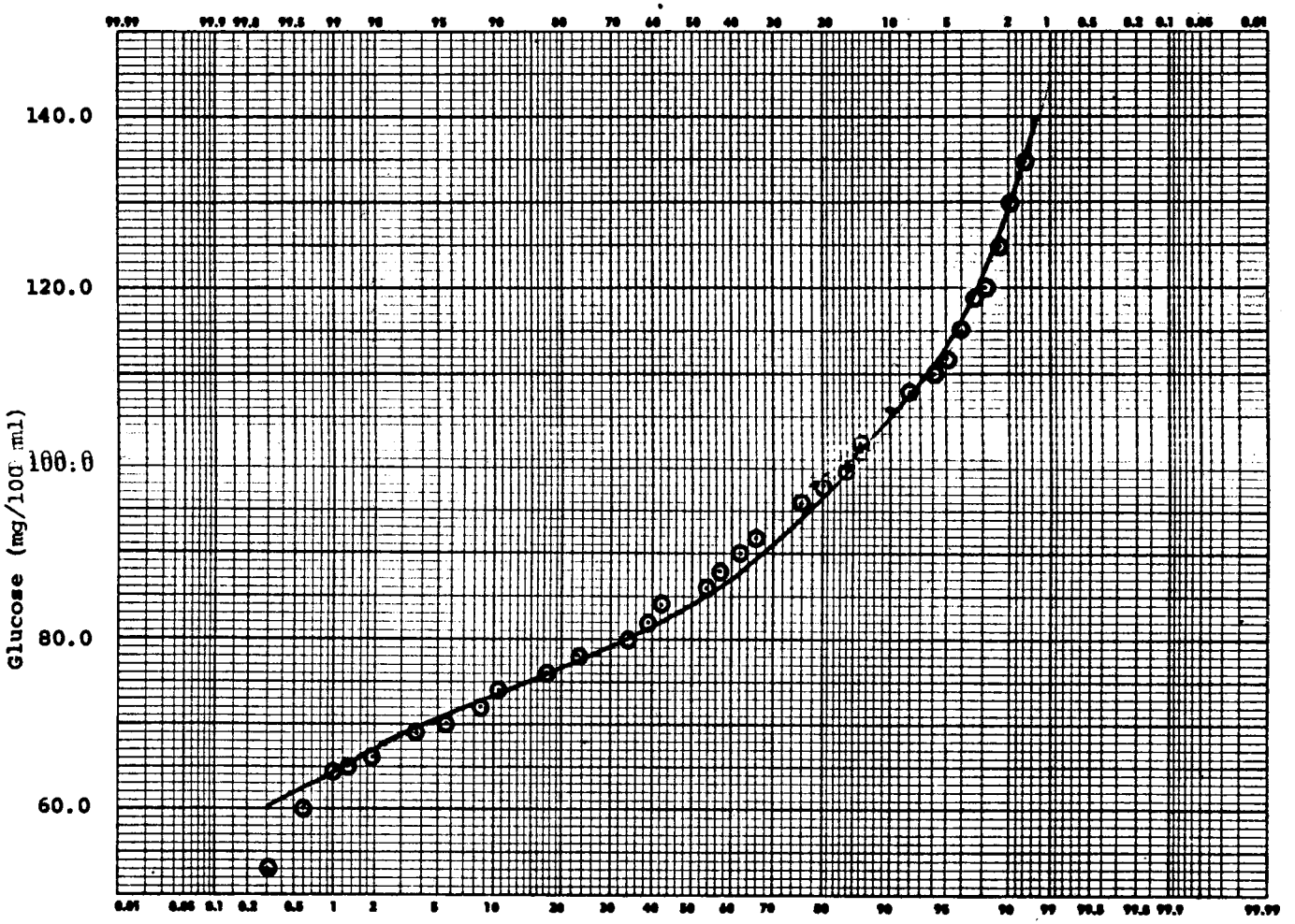


Figure 29. Glucose for Age Group 55-59

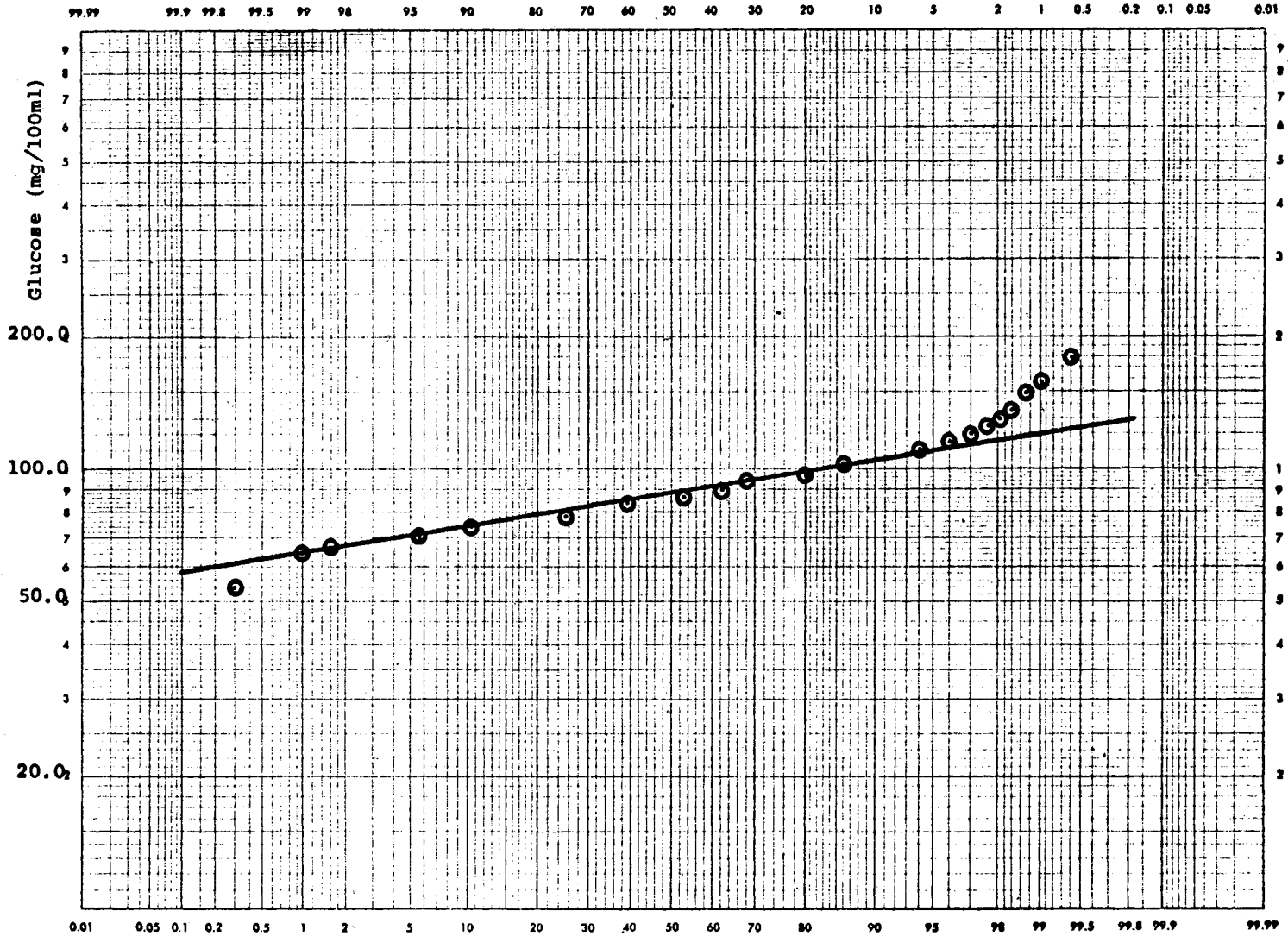


Figure 30. Glucose for Age Group 55-59

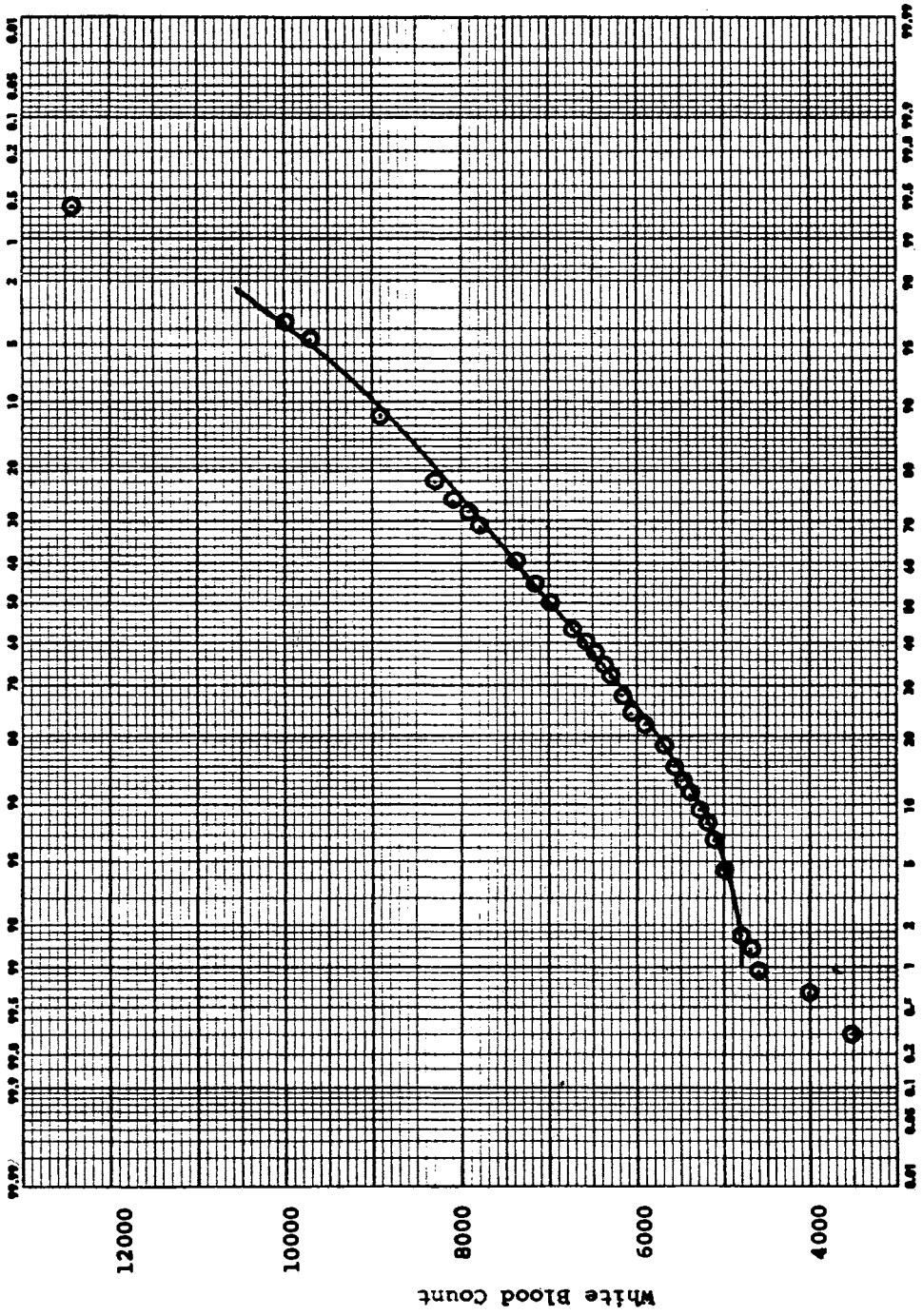


Figure 31. White Blood Count for Age Group 55-59

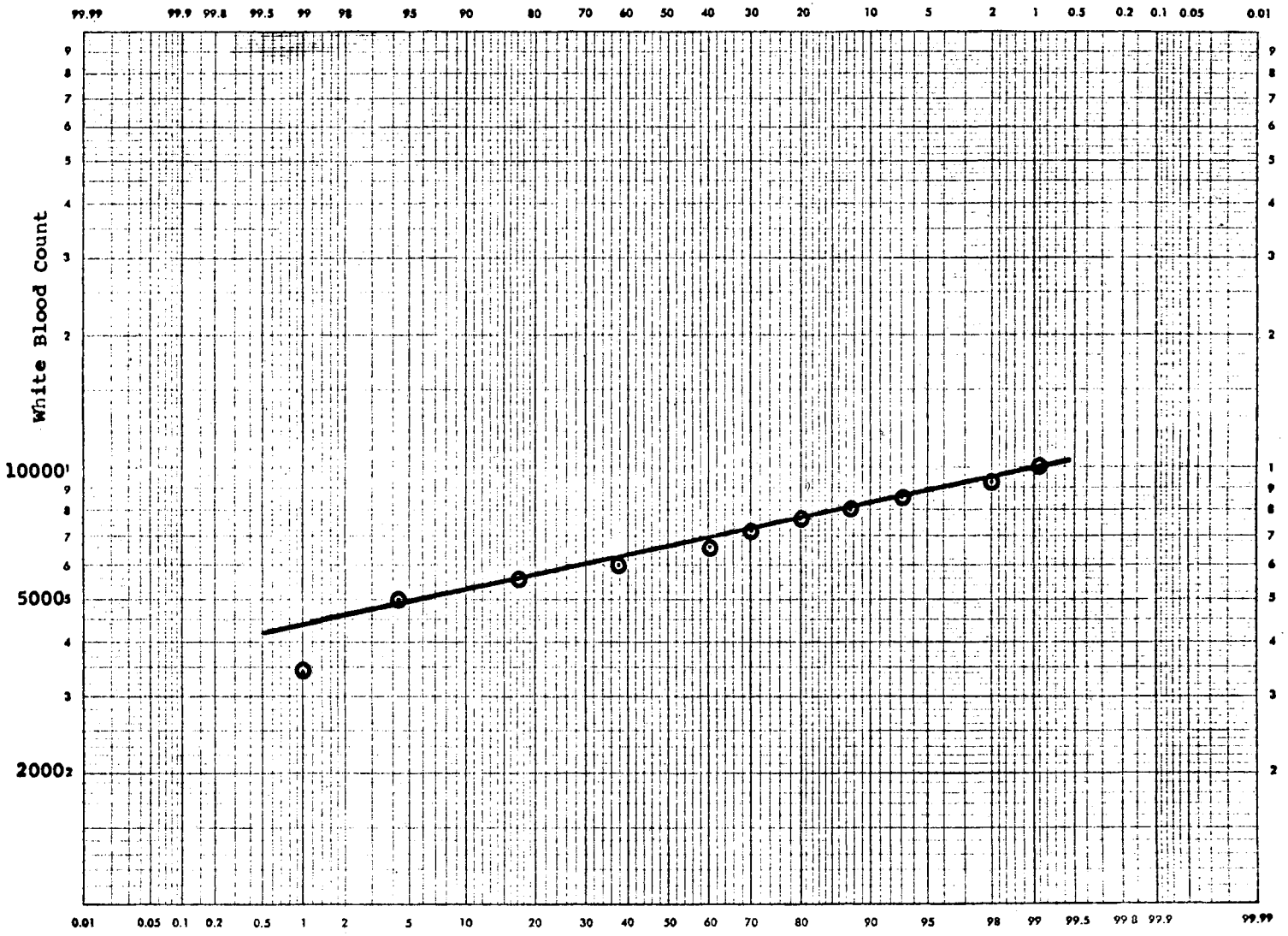


Figure 32. White Blood Count for Age Group 55-59

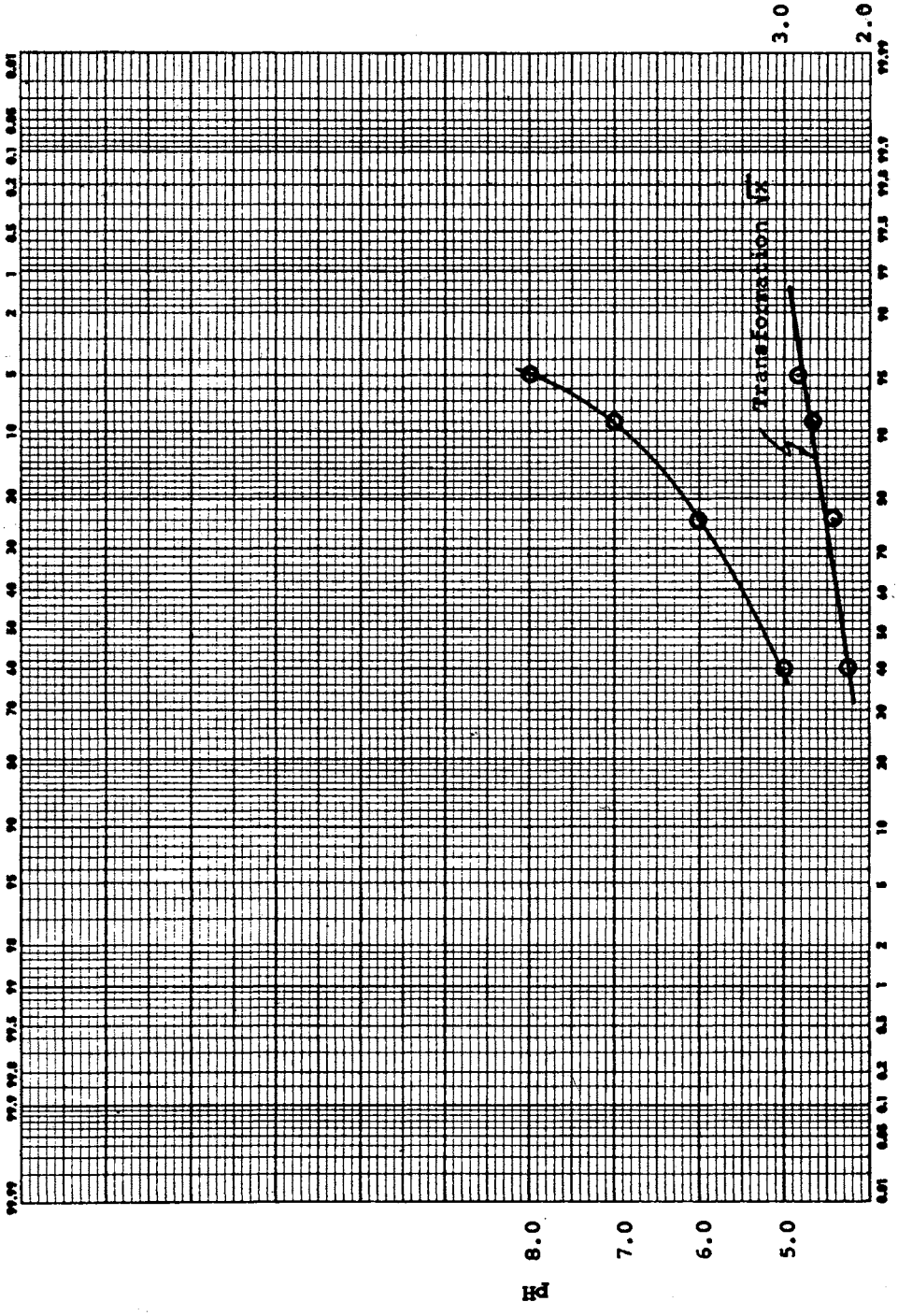


Figure 33. Urine pH for Age Group 55-59

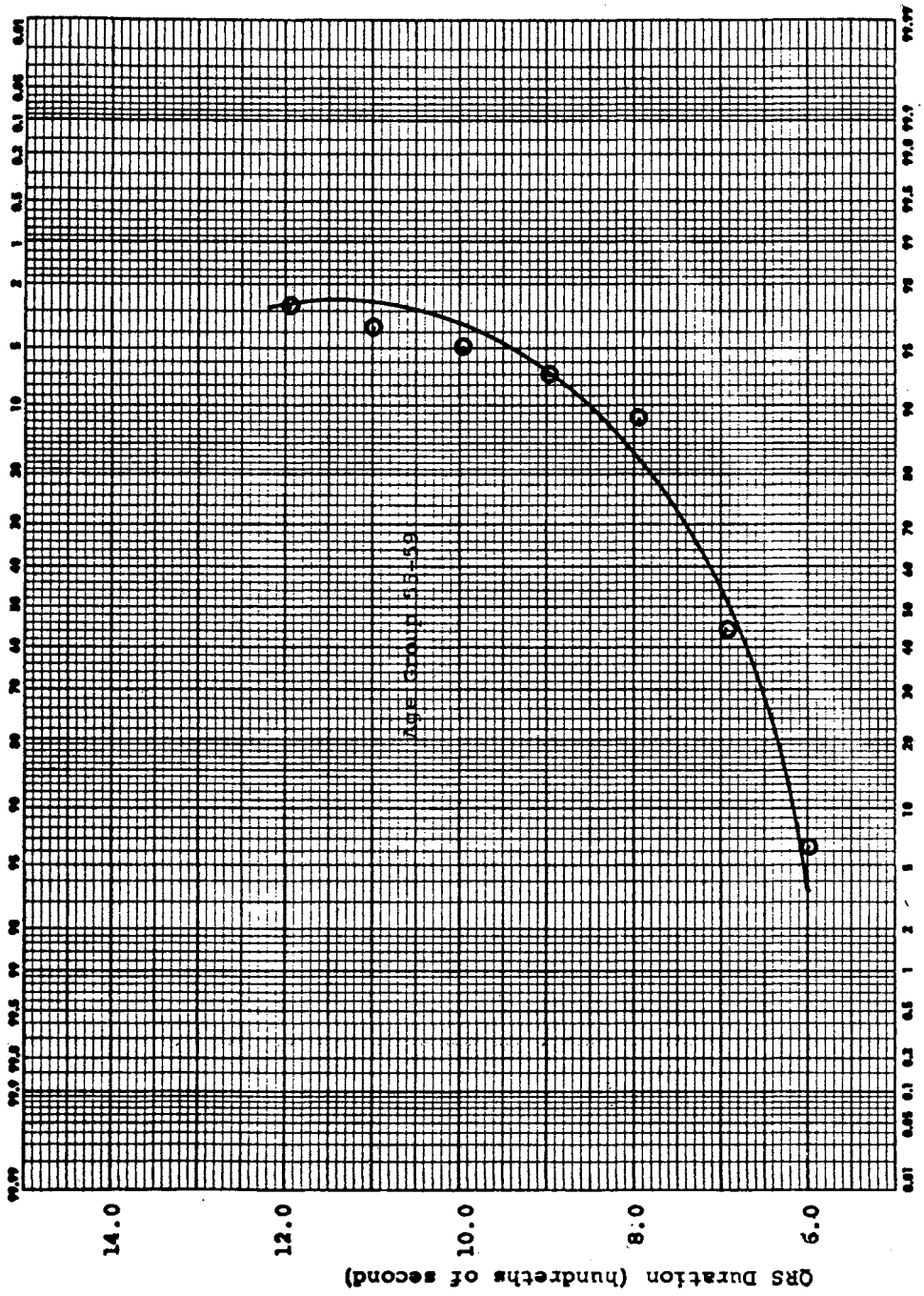


Figure 34. QRS Duration for Age Group 55-59

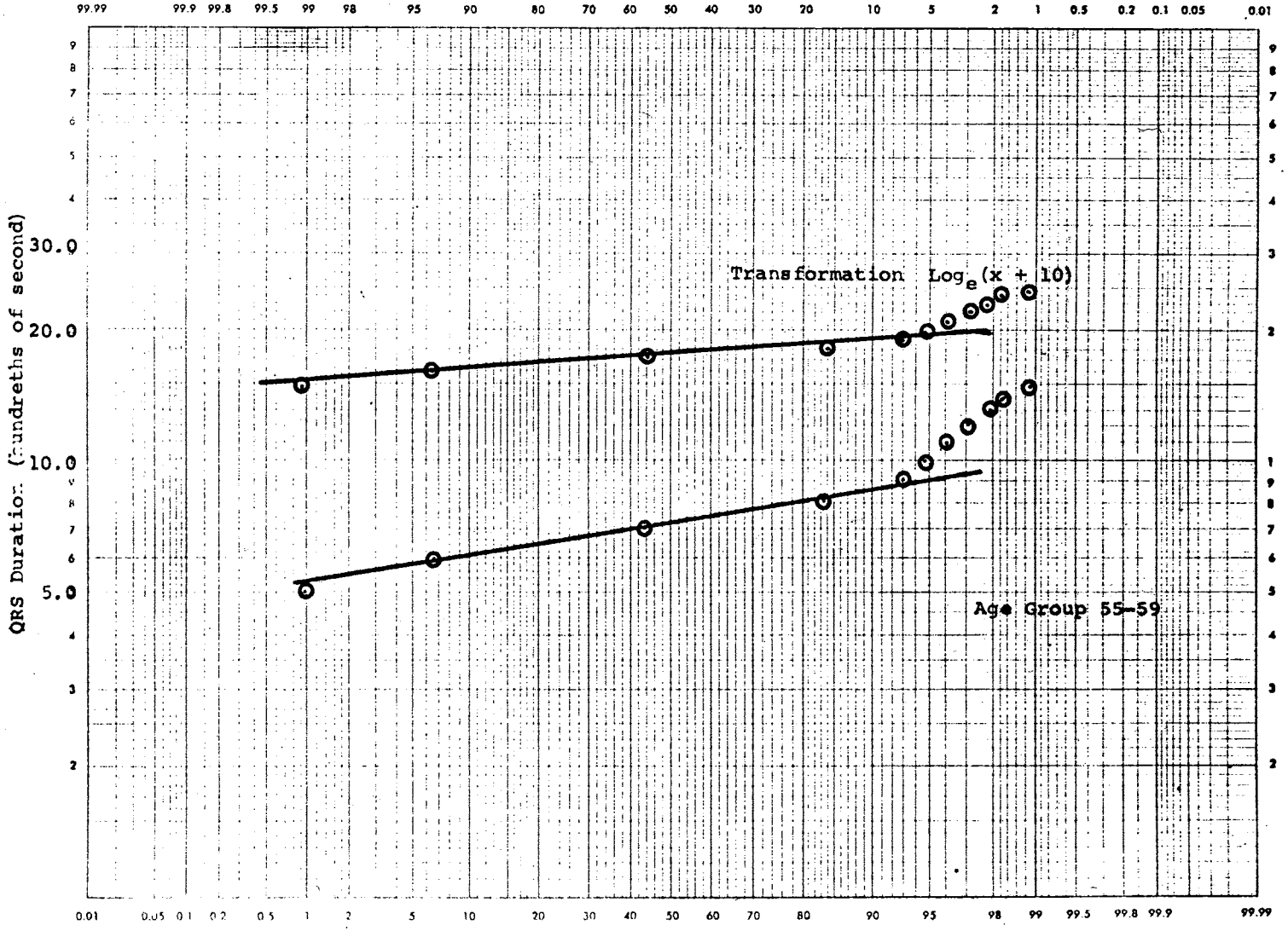


Figure 35. QRS Duration for Age Group 55-59

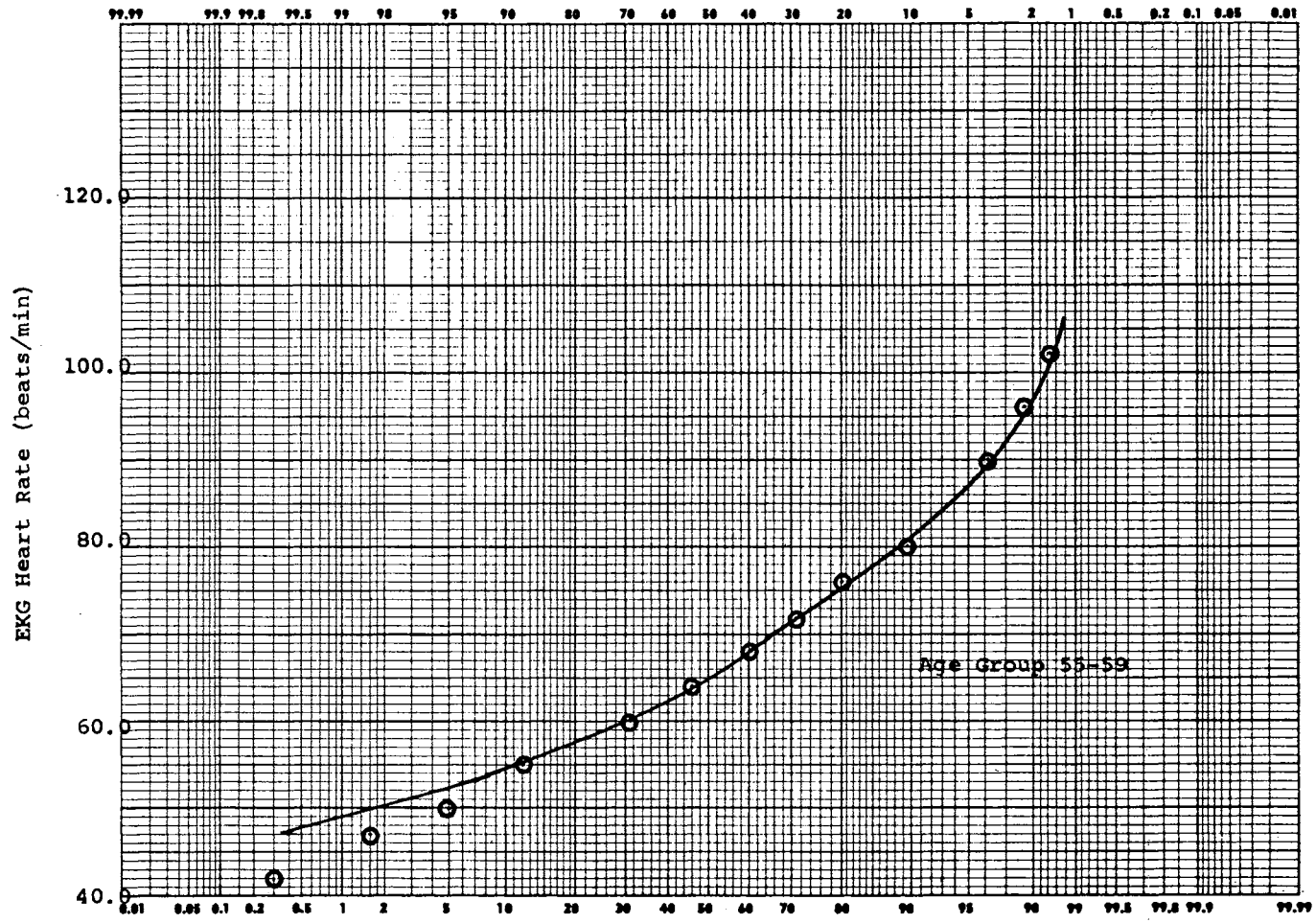


Figure 36. EKG Heart Rate for Age Group 55-59

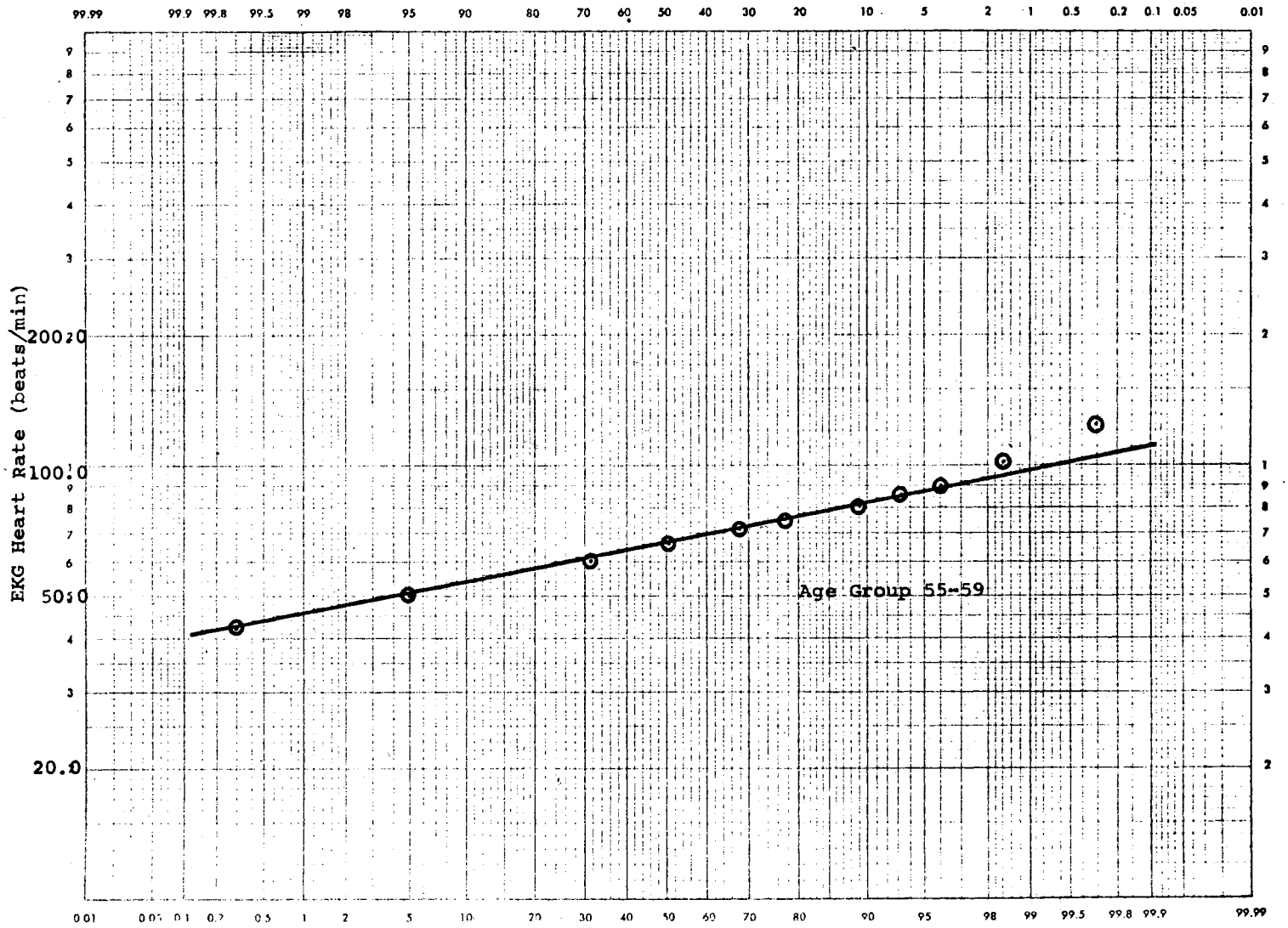


Figure 37. EKG Heart Rate for Age Group 55-59

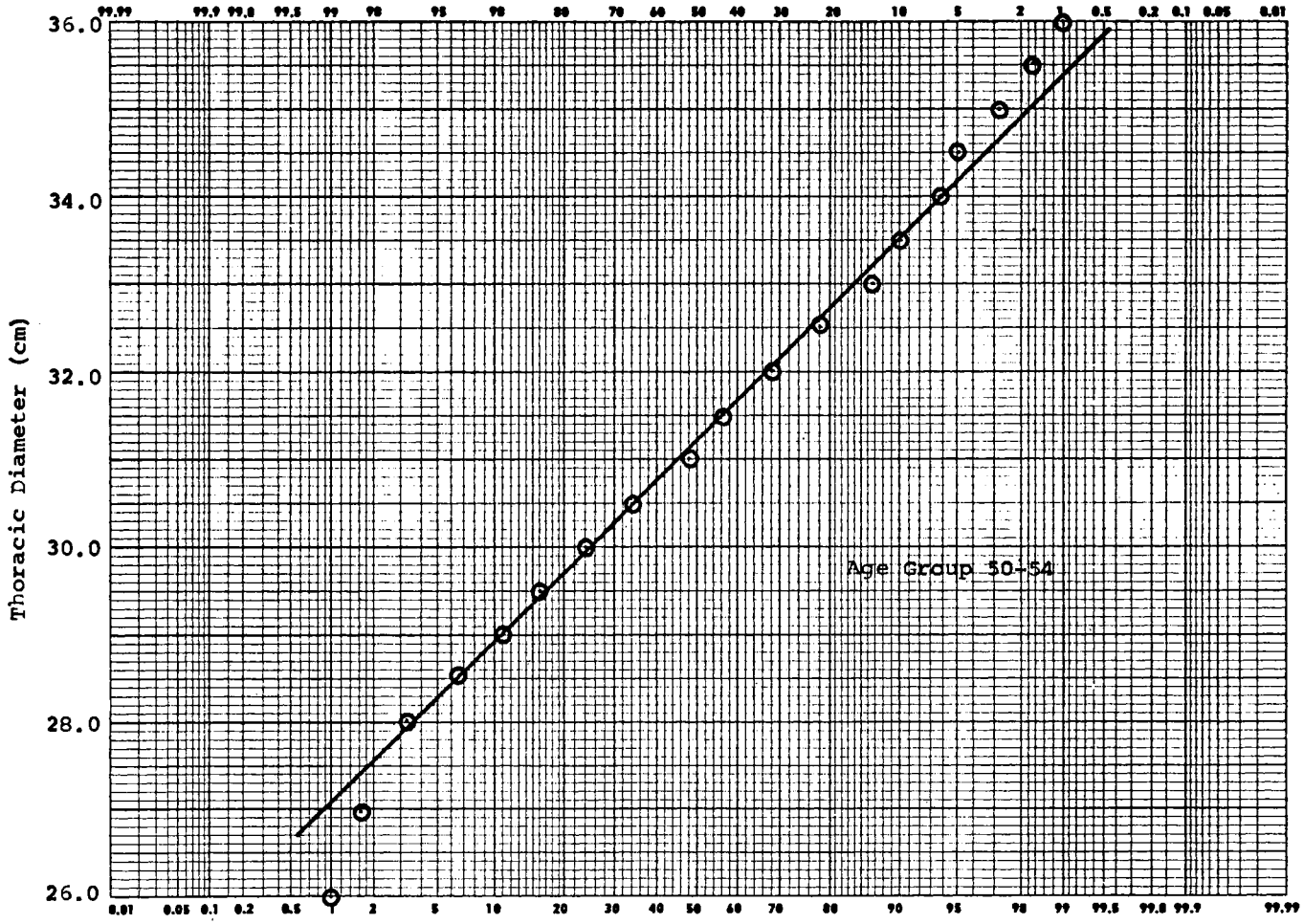


Figure 38. Thoracic Diameter for Age Group 50-54

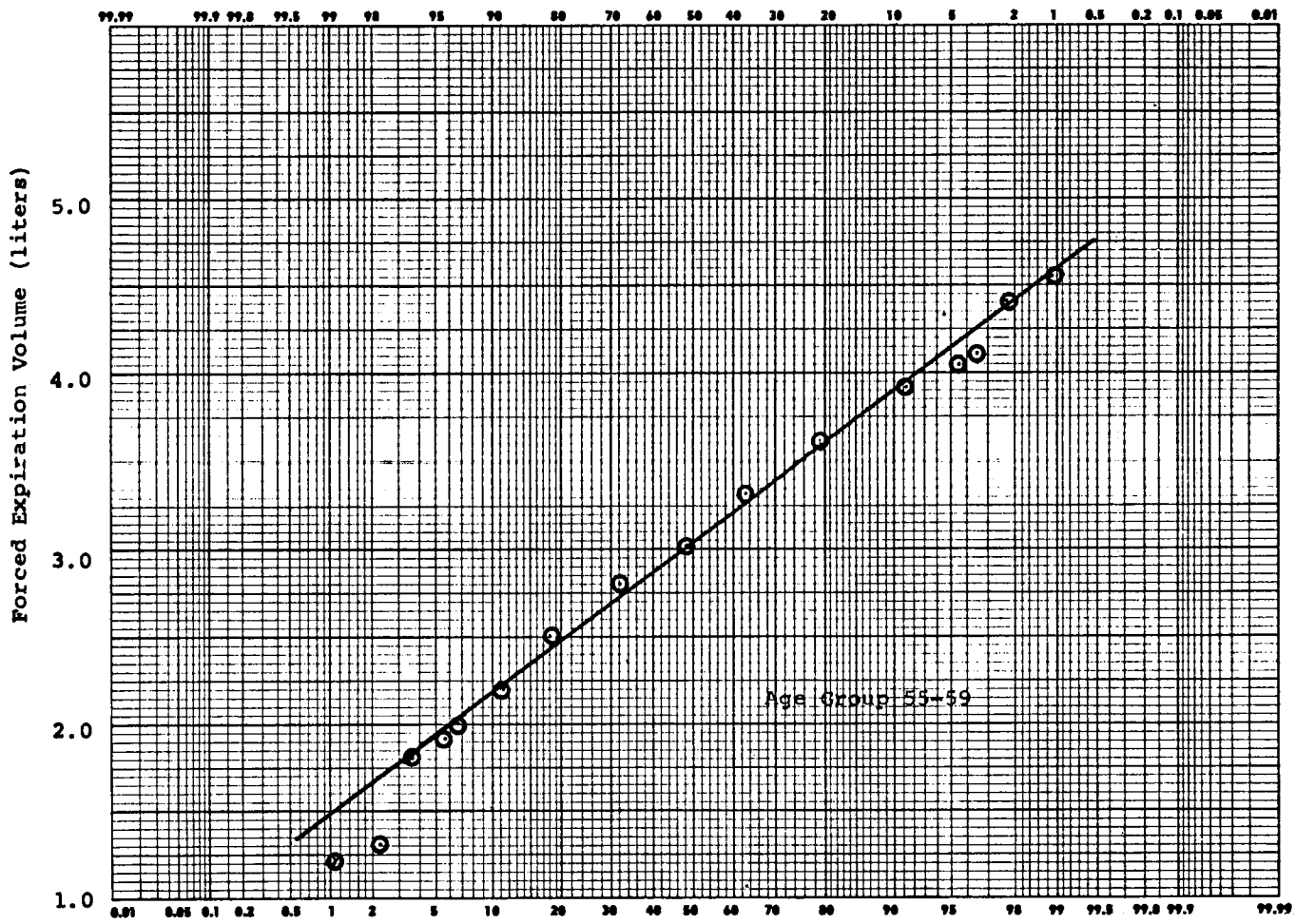


Figure 39. Forced Expiration Volume for Age Group 55-59

APPENDIX C

ESTIMATES OF POPULATION STATISTICS FOR EACH AGE GROUP

TABLE XXXVI

POPULATION STATISTICS FOR AGE GROUP 20-29

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 20 ***** MAXIMUM AGE = 29

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.78022	2.50245	91	.0892
2	WEIGHT	169.25000	24.16098	88	.0928
3	ACTUAL WEIGHT/IDEAL WEIGHT	.07202	.12251	87	.0796
4	PULSE	4.26977	.11613	89	.1446
5	SYSTOLIC BLOOD PRESSURE	4.71612	.11613	91	.0753
6	DIASTOLIC BLOOD PRESSURE	4.22953	.13218	92	.0926
7	ARM SKIN FOLDS	2.24030	.49922	47	.0756
8	BACK SKIN FOLDS	2.62575	.44109	45	.1253
9	HEMATOCRIT	45.04255	2.04737	94	.1158
10	WHITE BLOOD COUNT	4.16870	.17868	94	.0873
11	GLUCOSE FASTING	4.40375	.11779	94	.0627
12	CHOLESTEROL	5.32387	.18705	95	.0598
13	URIC ACID	1.60144	.18235	94	.0872
14	SGPT	2.48410	.55233	86	.1086
15	TOTAL BILIRUBIN	3.27518	1.73886	95	.1088
16	URINE RED CELL COUNT	1.03902	.28940	55	.0447
17	URINE WHITE CELL COUNT	2.46303	1.03436	54	.0668
18	URINE SPECIFIC GRAVITY	1.76458	.55421	96	.1181
19	URINE PH	2.44542	.14196	96	.0484
20	TOTAL HEART DIAMETER	12.42857	1.47196	49	.1345
21	THORACIC DIAMETER	30.52128	1.51786	47	.0846
22	TOTAL HEART DIA/THORACIC DIAMETER	.40702	.04330	47	.0852
23	EKG HEART RATE	4.14209	.14247	92	.0627
24	PR INTERVAL	2.74853	.11895	48	.0458
25	QRS INTERVAL	2.87081	.05041	49	.0334
26	QRS AXIS	57.86792	26.69849	53	.1615
27	VITAL CAPACITY (VC)	4.94091	.33758	22	.1134
28	FORCED EXPIRATION VOLUME (FEV1)	4.14545	.38635	22	.1359

TABLE XXXVII

POPULATION STATISTICS FOR AGE GROUP 30-34

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 30 ***** MAXIMUM AGE = 34

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	71.20952	2.33014	210	.1209
2	WEIGHT	178.95238	22.61129	210	.0610
3	ACTUAL WEIGHT/IDEAL WEIGHT	.11337	.11199	207	.0296
4	PULSE	4.28928	.12492	211	.0876
5	SYSTOLIC BLOOD PRESSURE	4.71243	.11426	212	.0628
6	DIASTOLIC BLOOD PRESSURE	4.24723	.13254	213	.0993
7	ARM SKIN FOLDS	2.37371	.48632	82	.1049
8	BACK SKIN FOLDS	2.74519	.39026	83	.0736
9	HEMATOCRIT	45.03211	2.31681	218	.0921
10	WHITE BLOOD COUNT	4.19629	.20028	220	.0670
11	GLUCOSE FASTING	4.43363	.12458	214	.0606
12	CHOLESTEROL	5.37918	.18012	215	.0660
13	URIC ACID	1.59138	.21779	209	.0964
14	SGPT	2.55272	.58502	195	.0664
15	TOTAL BILIRUBIN	2.93564	1.83824	217	.0873
16	URINE RED CELL COUNT	1.02746	.19610	103	.0461
17	URINE WHITE CELL COUNT	2.63159	1.63348	104	.0963
18	URINE SPECIFIC GRAVITY	1.68813	.61774	219	.1023
19	URINE PH	2.42510	.15346	220	.0471
20	TOTAL HEART DIAMETER	12.50595	1.08798	84	.1026
21	THORACIC DIAMETER	31.39412	1.45793	85	.0873
22	TOTAL HEART DIA/THORACIC DIAMETER	.39890	.02988	83	.0690
23	EKG HEART RATE	4.17502	.16460	211	.0700
24	PR INTERVAL	2.74187	.12296	84	.0328
25	QRS INTERVAL	2.86638	.04446	83	.0327
26	QRS AXIS	50.52083	29.31498	96	.1423
27	VITAL CAPACITY (VC)	5.00417	.72561	24	.1103
28	FORCED EXPIRATION VOLUME (FEV1)	4.06667	.60409	24	.1097

TABLE XXXVIII

POPULATION STATISTICS FOR AGE GROUP 35-39

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 35 ***** MAXIMUM AGE = 39

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	71.06677	2.37798	629	.0906
2	WEIGHT	180.16720	23.88601	628	.0543
3	ACTUAL WEIGHT/IDEAL WEIGHT	.12206	.11322	619	.0399
4	PULSE	4.28925	.12739	626	.0886
5	SYSTOLIC BLOOD PRESSURE	4.72552	.10998	631	.0650
6	DIASTOLIC BLOOD PRESSURE	4.27553	.12785	630	.0946
7	ARM SKIN FOLDS	2.34517	.47103	227	.0646
8	BACK SKIN FOLDS	2.79362	.38854	232	.0636
9	HEMATOCRIT	44.84555	2.26857	641	.0887
10	WHITE BLOOD COUNT	4.20398	.21945	639	.0651
11	GLUCOSE FASTING	4.42649	.14220	634	.0466
12	CHOLESTEROL	5.41515	.16587	635	.0506
13	URIC ACID	1.58283	.23071	631	.0512
14	SGPT	2.56182	.53696	584	.0785
15	TOTAL BILIRUBIN	3.29687	1.71927	641	.1396
16	URINE RED CELL COUNT	1.04123	.28384	304	.0463
17	URINE WHITE CELL COUNT	2.37021	1.25315	303	.0570
18	URINE SPECIFIC GRAVITY	1.68459	.59466	636	.0977
19	URINE PH	2.42942	.15451	641	.0504
20	TOTAL HEART DIAMETER	12.95902	1.37973	244	.0830
21	THORACIC DIAMETER	31.42008	1.76114	244	.0800
22	TOTAL HEART DIA/THORACIC DIAMETER	.41284	.03790	242	.0449
23	EKG HEART RATE	4.16002	.16025	625	.0616
24	PR INTERVAL	2.76394	.12351	244	.0277
25	QRS INTERVAL	2.86596	.04792	244	.0172
26	QRS AXIS	44.40614	29.99764	293	.1236
27	VITAL CAPACITY (VC)	4.77317	.69921	82	.0918
28	FORCED EXPIRATION VOLUME (FEV1)	3.84938	.55591	81	.0770

TABLE XXXIX

POPULATION STATISTICS FOR AGE GROUP 40-44

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 40 ***** MAXIMUM AGE = 44

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.63593	2.44742	629	.0991
2	WEIGHT	179.03840	23.55426	625	.0441
3	ACTUAL WEIGHT/IDEAL WEIGHT	.12798	.11182	617	.0365
4	PULSE	4.30512	.12888	616	.0925
5	SYSTOLIC BLOOD PRESSURE	4.74511	.12146	627	.0590
6	DIASTOLIC BLOOD PRESSURE	4.29312	.13166	627	.0801
7	ARM SKIN FOLDS	2.39286	.46950	209	.0952
8	BACK SKIN FOLDS	2.85728	.34640	208	.1003
9	HEMATOCRIT	45.02669	2.39646	637	.1001
10	WHITE BLOOD COUNT	4.26763	.20753	634	.0319
11	GLUCOSE FASTING	4.44959	.15520	632	.0477
12	CHOLESTEROL	5.44792	.17393	631	.0508
13	URIC ACID	1.59886	.22532	629	.0527
14	SGPT	2.52611	.58578	588	.0645
15	TOTAL BILIRUBIN	3.33350	1.76472	641	.1330
16	URINE RED CELL COUNT	1.00000	.00000	291	.0000
17	URINE WHITE CELL COUNT	2.27799	1.15144	290	.0675
18	URINE SPECIFIC GRAVITY	1.61717	.60558	635	.0823
19	URINE PH	2.42919	.15796	641	.0466
20	TOTAL HEART DIAMETER	13.36062	1.40927	226	.0818
21	THORACIC DIAMETER	31.54867	1.82204	226	.0985
22	TOTAL HEART DIA/THORACIC DIAMETER	.42441	.03831	225	.0323
23	EKG HEART RATE	4.18496	.16644	624	.0663
24	PR INTERVAL	2.76649	.11640	220	.0318
25	QRS INTERVAL	2.86067	.05043	222	.0269
26	QRS AXIS	40.00000	30.43699	274	.0935
27	VITAL CAPACITY (VC)	4.60660	.62374	93	.0755
28	FORCED EXPIRATION VOLUME (FEV1)	3.63478	.50913	92	.1111

TABLE XL

POPULATION STATISTICS FOR AGE GROUP 45-49

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 45 ***** MAXIMUM AGE = 49

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.37292	2.32552	783	.0948
2	WEIGHT	179.02561	24.08360	781	.0325
3	ACTUAL WEIGHT/IDEAL WEIGHT	.13552	.11908	775	.0384
4	PULSE	4.30483	.12616	768	.0899
5	SYSTOLIC BLOOD PRESSURE	4.76393	.13515	781	.0749
6	DIASTOLIC BLOOD PRESSURE	4.31823	.14214	780	.0814
7	ARM SKIN FOLDS	2.32179	.48322	416	.0765
8	BACK SKIN FOLDS	2.84030	.38621	410	.0919
9	HEMATOCRIT	45.12830	2.50664	795	.0870
10	WHITE BLOOD COUNT	4.24244	.20558	792	.0471
11	GLUCOSE FASTING	4.43193	.14189	786	.0446
12	CHOLESTEROL	5.46495	.17113	791	.0386
13	URIC ACID	1.58227	.21939	793	.0512
14	SGPT	2.45743	.53722	716	.0951
15	TOTAL BILIRUBIN	3.36532	1.74568	801	.1225
16	URINE RED CELL COUNT	1.02742	.22638	514	.0404
17	URINE WHITE CELL COUNT	2.29459	1.23541	512	.0563
18	URINE SPECIFIC GRAVITY	1.59214	.59560	789	.0811
19	URINE PH	2.42281	.16659	797	.0462
20	TOTAL HEART DIAMETER	13.36885	1.31009	427	.1006
21	THORACIC DIAMETER	31.24825	1.73361	429	.0710
22	TOTAL HEART DIA/THORACIC DIAMETER	.42805	.03648	425	.0364
23	EKG HEART RATE	4.18900	.15774	783	.0498
24	PR INTERVAL	2.78773	.12904	430	.0225
25	QRS INTERVAL	2.86084	.05213	426	.0186
26	QRS AXIS	34.62800	31.33505	500	.0903
27	VITAL CAPACITY (VC)	4.41532	.65927	235	.0535
28	FORCED EXPIRATION VOLUME (FEV1)	3.43248	.56338	234	.0813

TABLE XLI

POPULATION STATISTICS FOR AGE GROUP 50-54

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ••••• SEX • MALE ••••• MINIMUM AGE • 50 ••••• MAXIMUM AGE • 54

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.41586	2.35577	618	.1084
2	WEIGHT	178.75527	23.90104	617	.0507
3	ACTUAL WEIGHT/IDEAL WEIGHT	.13405	.11910	614	.0321
4	PULSE	4.30551	.12449	611	.0885
5	SYSTOLIC BLOOD PRESSURE	4.78679	.14029	621	.0777
6	DIASTOLIC BLOOD PRESSURE	4.32911	.14076	622	.0994
7	ARM SKIN FOLDS	2.32403	.48084	336	.0757
8	BACK SKIN FOLDS	2.80469	.39137	333	.0645
9	HEMATOCRIT	45.29630	2.52071	621	.1014
10	WHITE BLOOD COUNT	4.26281	.21950	619	.0410
11	GLUCOSE FASTING	4.45936	.14535	623	.0552
12	CHOLESTEROL	5.47241	.16669	623	.0405
13	URIC ACID	1.57758	.22130	624	.0537
14	SGPT	2.43885	.52595	542	.0884
15	TOTAL BILIRUBIN	3.36234	1.75568	631	.1372
16	URINE RED CELL COUNT	1.04050	.23616	419	.0538
17	URINE WHITE CELL COUNT	2.32366	1.20245	418	.0602
18	URINE SPECIFIC GRAVITY	1.61694	.58474	620	.0900
19	URINE PH	2.41331	.16162	624	.0500
20	TOTAL HEART DIAMETER	13.50588	1.33535	340	.1111
21	THORACIC DIAMETER	31.44412	1.76657	340	.0714
22	TOTAL HEART DIA/THORACIC DIAMETER	.43013	.03925	338	.0351
23	EKG HEART RATE	4.19655	.15891	624	.0535
24	PR INTERVAL	2.79407	.13103	344	.0231
25	QRS INTERVAL	2.86536	.05510	340	.0189
26	QRS AXIS	30.75366	32.25161	410	.0908
27	VITAL CAPACITY (VC)	4.36259	.66767	147	.0531
28	FORCED EXPIRATION VOLUME (FEV1)	3.35442	.58722	147	.0936

TABLE XLII

POPULATION STATISTICS FOR AGE GROUP 55-59

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX • MALE ***** MINIMUM AGE • 55 ***** MAXIMUM AGE • 59

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.18874	2.35563	302	.1133
2	WEIGHT	179.71617	23.51691	303	.0432
3	ACTUAL WEIGHT/IDEAL WEIGHT	.14808	.11484	299	.0595
4	PULSE	4.30644	.13021	292	.0891
5	SYSTOLIC BLOOD PRESSURE	4.83069	.14870	300	.0761
6	DIASTOLIC BLOOD PRESSURE	4.35079	.13440	300	.0736
7	ARM SKIN FOLDS	2.36593	.44924	163	.0739
8	BACK SKIN FOLDS	2.85475	.32726	159	.0917
9	HEMATOCRIT	45.16938	2.47550	307	.0864
10	WHITE BLOOD COUNT	4.25646	.19775	306	.0389
11	GLUCOSE FASTING	4.46673	.14265	301	.0743
12	CHOLESTEROL	5.47004	.16590	306	.0363
13	URIC ACID	1.59338	.21947	305	.0539
14	SGPT	2.40551	.47473	267	.0819
15	TOTAL BILIRUBIN	3.38322	1.70924	311	.1610
16	URINE RED CELL COUNT	1.01421	.14142	199	.0350
17	URINE WHITE CELL COUNT	2.41437	1.20385	197	.0718
18	URINE SPECIFIC GRAVITY	1.57947	.59369	302	.1204
19	URINE PH	2.41265	.16854	305	.0471
20	TOTAL HEART DIAMETER	13.67160	1.29450	169	.1116
21	THORACIC DIAMETER	31.53254	1.76957	169	.0782
22	TOTAL HEART DIA/THORACIC DIAMETER	.43325	.03628	167	.0516
23	EKG HEART RATE	4.19813	.16336	303	.0569
24	PR INTERVAL	2.79676	.12606	171	.0266
25	QRS INTERVAL	2.87372	.06931	170	.0718
26	QRS AXIS	28.48705	32.55014	193	.0752
27	VITAL CAPACITY (VC)	4.04337	.60932	83	.0788
28	FORCED EXPIRATION VOLUME (FEV1)	3.10488	.57989	82	.0801

TABLE XLIII

POPULATION STATISTICS FOR AGE GROUP 60-64

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 60 ***** MAXIMUM AGE = 64

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	69.69388	2.31911	147	.0961
2	WEIGHT	178.04054	23.07983	148	.0413
3	ACTUAL WEIGHT/IDEAL WEIGHT	.15007	.11409	147	.0533
4	PULSE	4.30842	.11953	140	.1210
5	SYSTOLIC BLOOD PRESSURE	4.85058	.14078	145	.0548
6	DIASTOLIC BLOOD PRESSURE	4.34389	.13291	146	.0932
7	ARM SKIN FOLDS	2.25619	.45550	69	.1208
8	BACK SKIN FOLDS	2.79955	.40139	68	.1055
9	HEMATOCRIT	44.90476	2.19693	147	.1037
10	WHITE BLOOD COUNT	4.25209	.19590	147	.0793
11	GLUCOSE FASTING	4.49220	.17354	145	.0852
12	CHOLESTEROL	5.46893	.15984	148	.0717
13	URIC ACID	1.59746	.19744	146	.0970
14	SGPT	2.37415	.46077	128	.0752
15	TOTAL BILIRUBIN	3.77102	1.58909	149	.1336
16	URINE RED CELL COUNT	1.12035	.46882	77	.0693
17	URINE WHITE CELL COUNT	2.61575	1.19149	77	.0535
18	URINE SPECIFIC GRAVITY	1.56892	.60943	148	.1194
19	URINE PH	2.43930	.16230	148	.0458
20	TOTAL HEART DIAMETER	13.55714	1.46343	70	.1065
21	THORACIC DIAMETER	31.14286	1.99118	70	.1064
22	TOTAL HEART DIA/THORACIC DIAMETER	.43556	.03636	69	.0681
23	EKG HEART RATE	4.17398	.14497	147	.0686
24	PR INTERVAL	2.79590	.13739	68	.0326
25	QRS INTERVAL	2.87575	.07000	69	.0530
26	QRS AXIS	20.94667	31.85520	75	.0741
27	VITAL CAPACITY (VC)	3.80541	.66872	37	.0929
28	FORCED EXPIRATION VOLUME (FEV1)	2.90811	.62110	37	.0823

TABLE XLIV

POPULATION STATISTICS FOR AGE GROUP 45-54

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 45 ***** MAXIMUM AGE = 54

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.39186	2.33817	1401	.1012
2	WEIGHT	178.90629	23.99500	1398	.0393
3	ACTUAL WEIGHT/IDEAL WEIGHT	.13487	.11905	1389	.0344
4	PULSE	4.30513	.12538	1379	.0882
5	SYSTOLIC BLOOD PRESSURE	4.77405	.13787	1402	.0772
6	DIASTOLIC BLOOD PRESSURE	4.32306	.14158	1402	.0872
7	ARM SKIN FOLDS	2.32279	.48184	752	.0769
8	BACK SKIN FOLDS	2.82434	.38867	743	.0803
9	HEMATOCRIT	45.20198	2.51332	1416	.0938
10	WHITE BLOOD COUNT	4.25138	.21196	1411	.0440
11	GLUCOSE FASTING	4.44406	.14403	1409	.0447
12	CHOLESTEROL	5.46824	.16917	1414	.0374
13	URIC ACID	1.58020	.22017	1417	.0476
14	SGPT	2.44942	.53226	1258	.0927
15	TOTAL BILIRUBIN	3.36400	1.74948	1432	.1290
16	URINE RED CELL COUNT	1.03330	.23079	933	.0467
17	URINE WHITE CELL COUNT	2.30766	1.22013	930	.0589
18	URINE SPECIFIC GRAVITY	1.60305	.59077	1409	.0854
19	URINE PH	2.41864	.16444	1421	.0476
20	TOTAL HEART DIAMETER	13.42960	1.32223	767	.1059
21	THORACIC DIAMETER	31.33485	1.74982	769	.0644
22	TOTAL HEART DIA/THORACIC DIAMETER	.42697	.03772	763	.0298
23	EKG HEART RATE	4.19235	.15825	1407	.0496
24	PR INTERVAL	2.79055	.12988	774	.0144
25	QRS INTERVAL	2.86284	.05348	766	.0183
26	QRS AXIS	32.68242	31.79227	910	.0875
27	VITAL CAPACITY (VC)	4.39503	.66214	382	.0461
28	FORCED EXPIRATION VOLUME (FEV1)	3.40236	.57319	381	.0847

TABLE XLV

POPULATION STATISTICS FOR AGE GROUP 55-64

ESTIMATES OF POPULATION STATISTICS

CLASSIFICATION ***** SEX = MALE ***** MINIMUM AGE = 55 ***** MAXIMUM AGE = 64

VARIABLE NUMBER	VARIABLE NAME	MEAN	STANDARD DEVIATION	SAMPLE SIZE	MAXIMUM DIFFERENCE (TEST FOR NORMALITY)
1	HEIGHT	70.02673	2.35266	449	.0960
2	WEIGHT	179.16630	23.36201	451	.0384
3	ACTUAL WEIGHT/IDEAL WEIGHT	.14873	.11447	446	.0436
4	PULSE	4.30708	.12671	432	.0855
5	SYSTOLIC BLOOD PRESSURE	4.83717	.14631	445	.0672
6	DIASTOLIC BLOOD PRESSURE	4.34853	.13381	446	.0798
7	ARM SKIN FOLDS	2.33329	.45292	232	.0903
8	BACK SKIN FOLDS	2.83822	.35111	227	.0923
9	HEMATOCRIT	45.08370	2.38965	454	.0904
10	WHITE BLOOD COUNT	4.25504	.19695	453	.0521
11	GLUCOSE FASTING	4.47501	.15364	446	.0613
12	CHOLESTEROL	5.46968	.16377	454	.0358
13	URIC ACID	1.59470	.21238	451	.0458
14	SGPT	2.39535	.46990	395	.0815
15	TOTAL BILIRUBIN	3.50883	1.67939	460	.1537
16	URINE RED CELL COUNT	1.04382	.27824	276	.0499
17	URINE WHITE CELL COUNT	2.47097	1.20164	274	.0605
18	URINE SPECIFIC GRAVITY	1.57600	.59825	450	.1044
19	URINE PH	2.42136	.16682	453	.0462
20	TOTAL HEART DIAMETER	13.63808	1.34406	239	.0992
21	THORACIC DIAMETER	31.41841	1.84158	239	.0885
22	TOTAL HEART DIA/THORACIC DIAMETER	.43393	.03625	236	.0507
23	EKG HEART RATE	4.19024	.15783	450	.0542
24	PR INTERVAL	2.79652	.12909	239	.0223
25	QRS INTERVAL	2.67431	.06937	239	.0668
26	QRS AXIS	26.37687	32.47522	268	.0723
27	VITAL CAPACITY (VC)	3.97000	.63505	120	.0560
28	FORCED EXPIRATION VOLUME (FEV1)	3.04370	.59740	119	.0579

APPENDIX D

GAMMA PROBABILITY PLOTS FOR EACH AGE GROUP

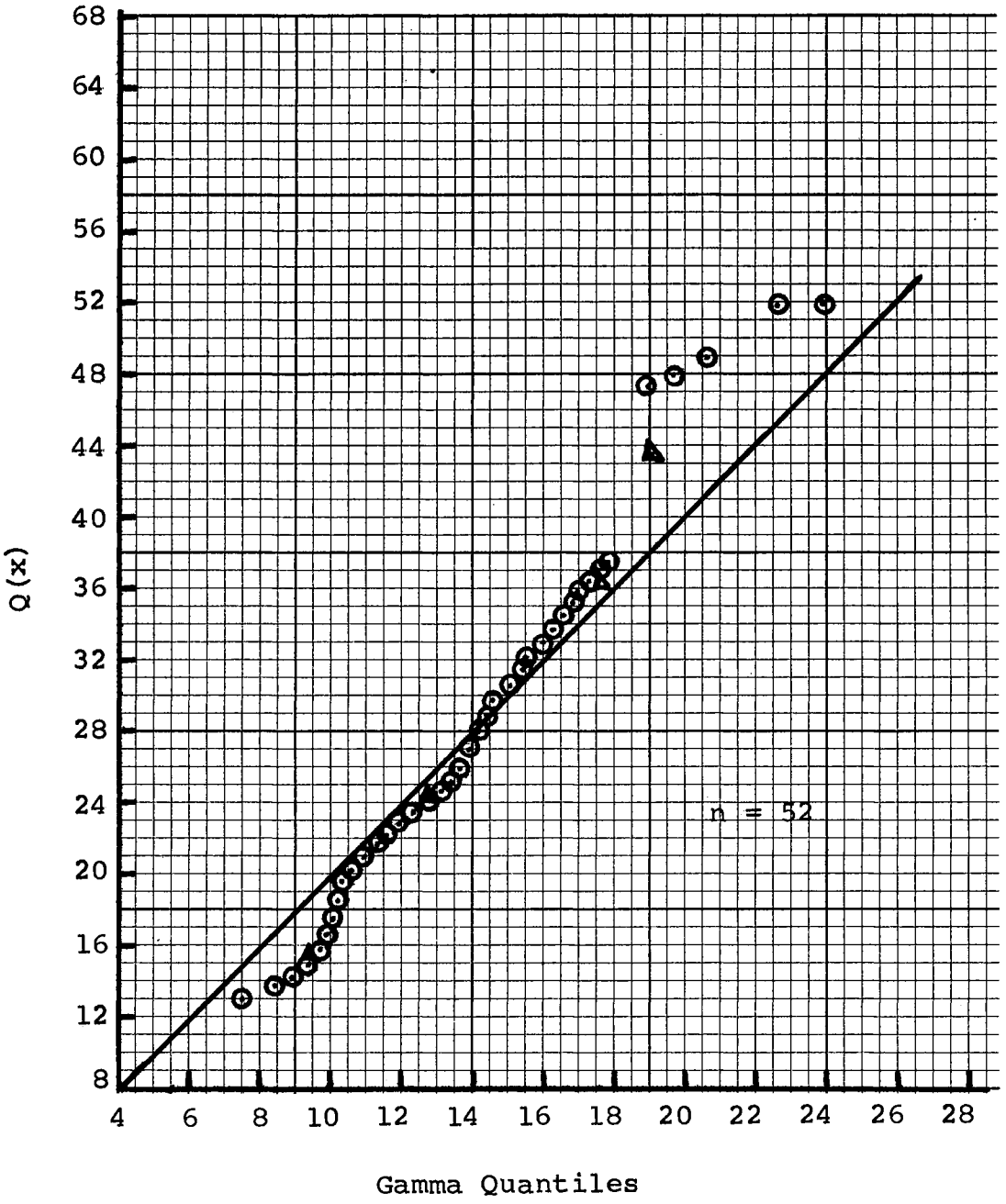


Figure 40. Multivariate Check for Age Group 20-29

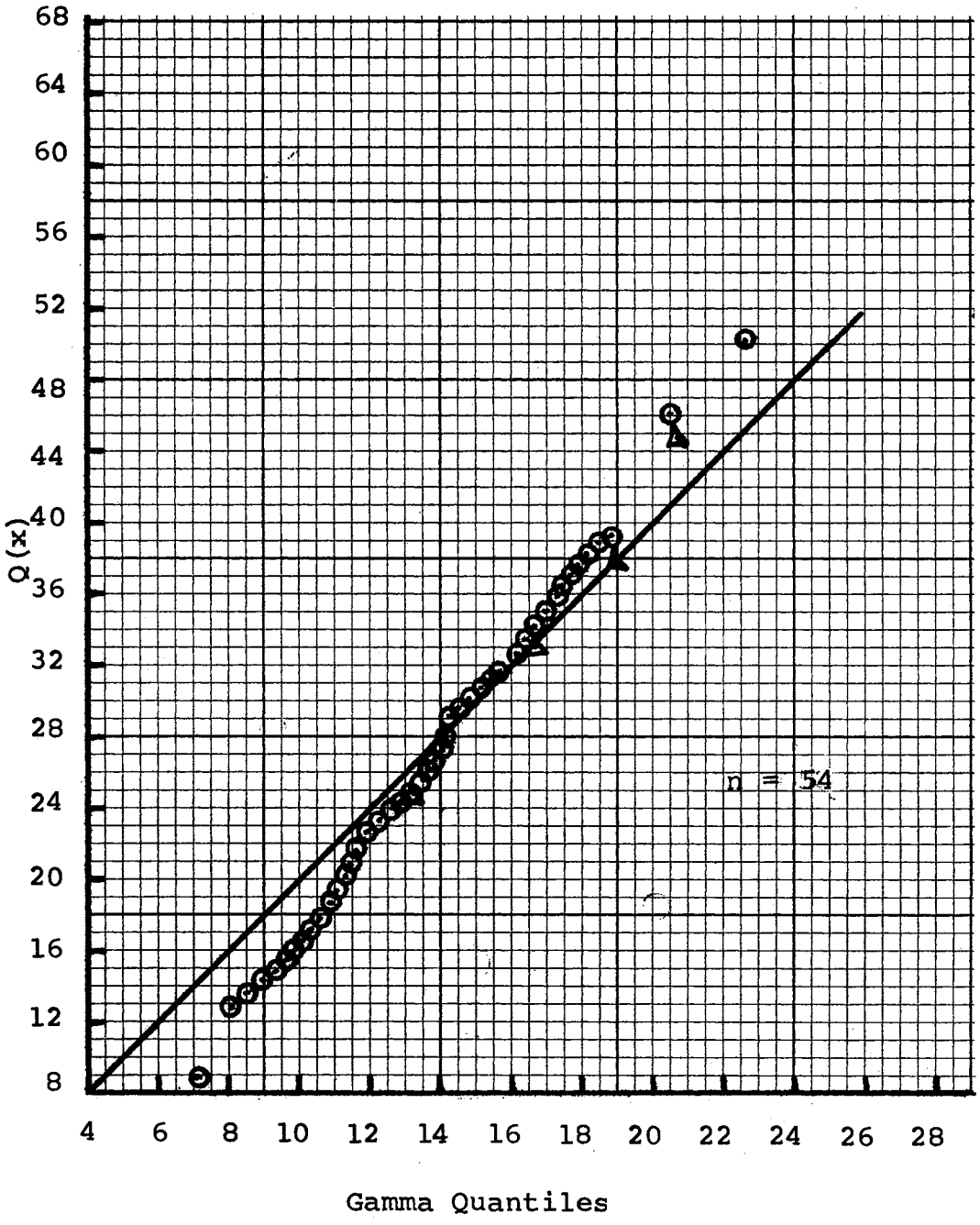


Figure 41. Multivariate Check for Age Group 30-34

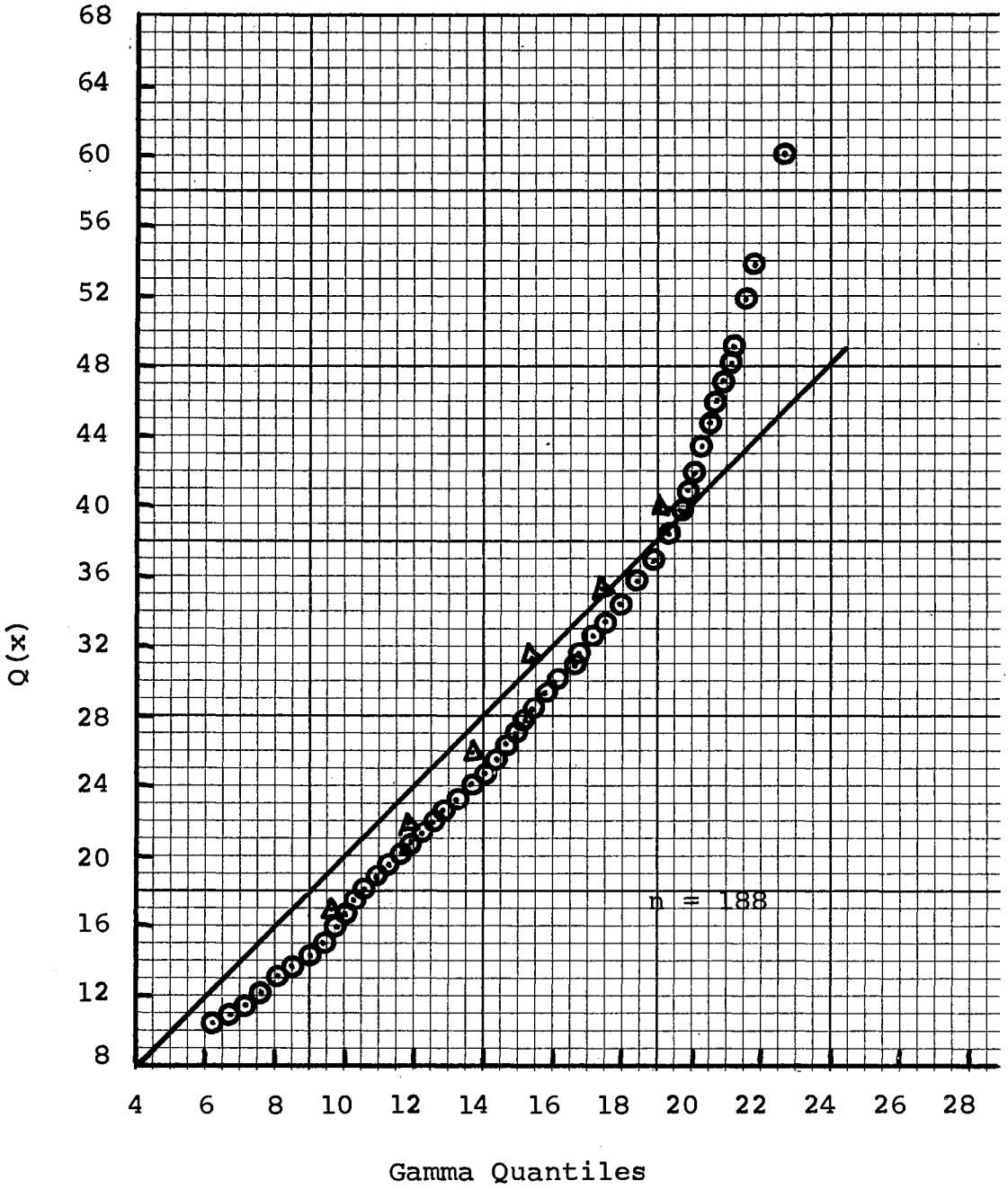


Figure 42. Multivariate Check for Age Group 35-39

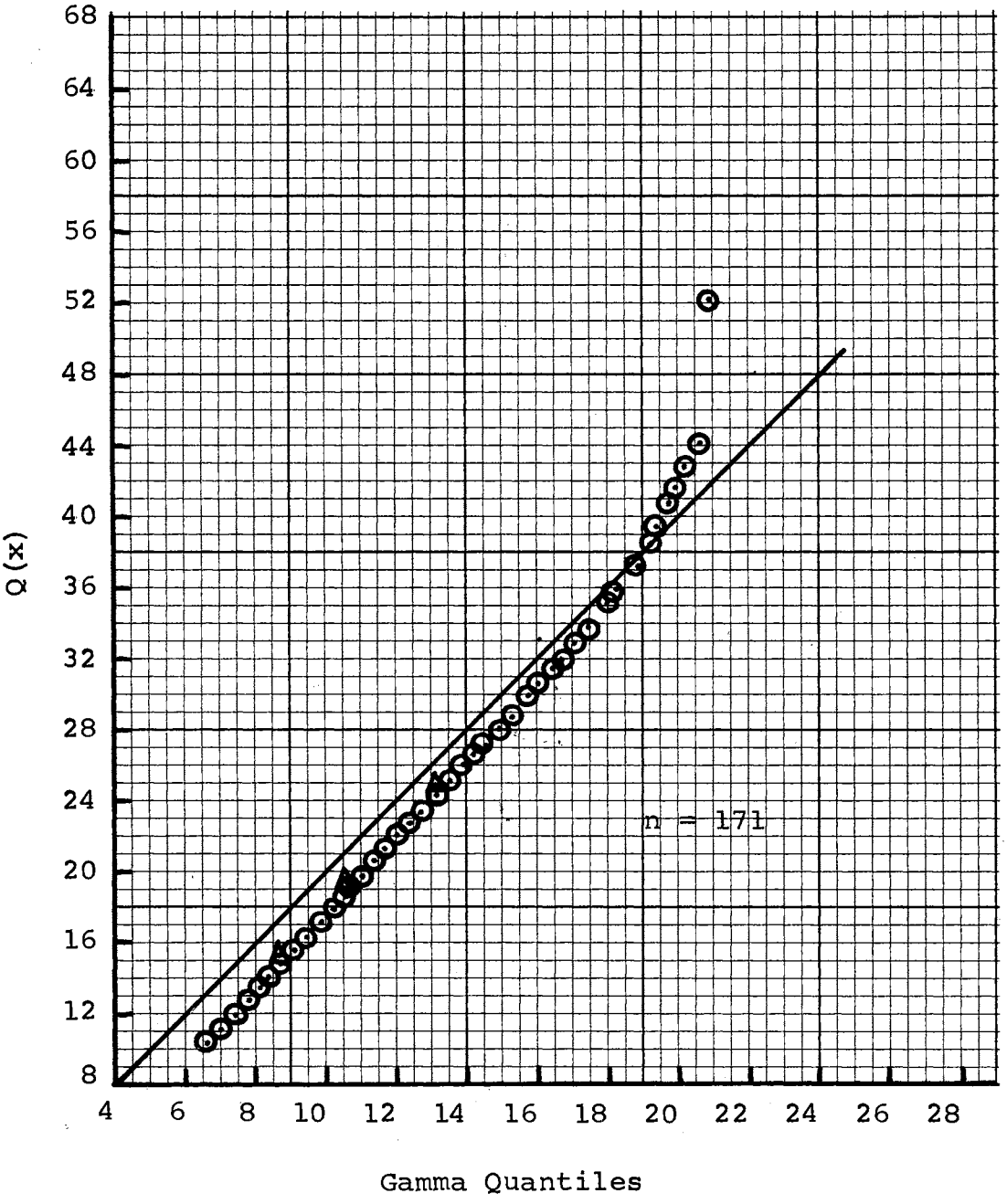


Figure 43. Multivariate Check for Age Group 40-44

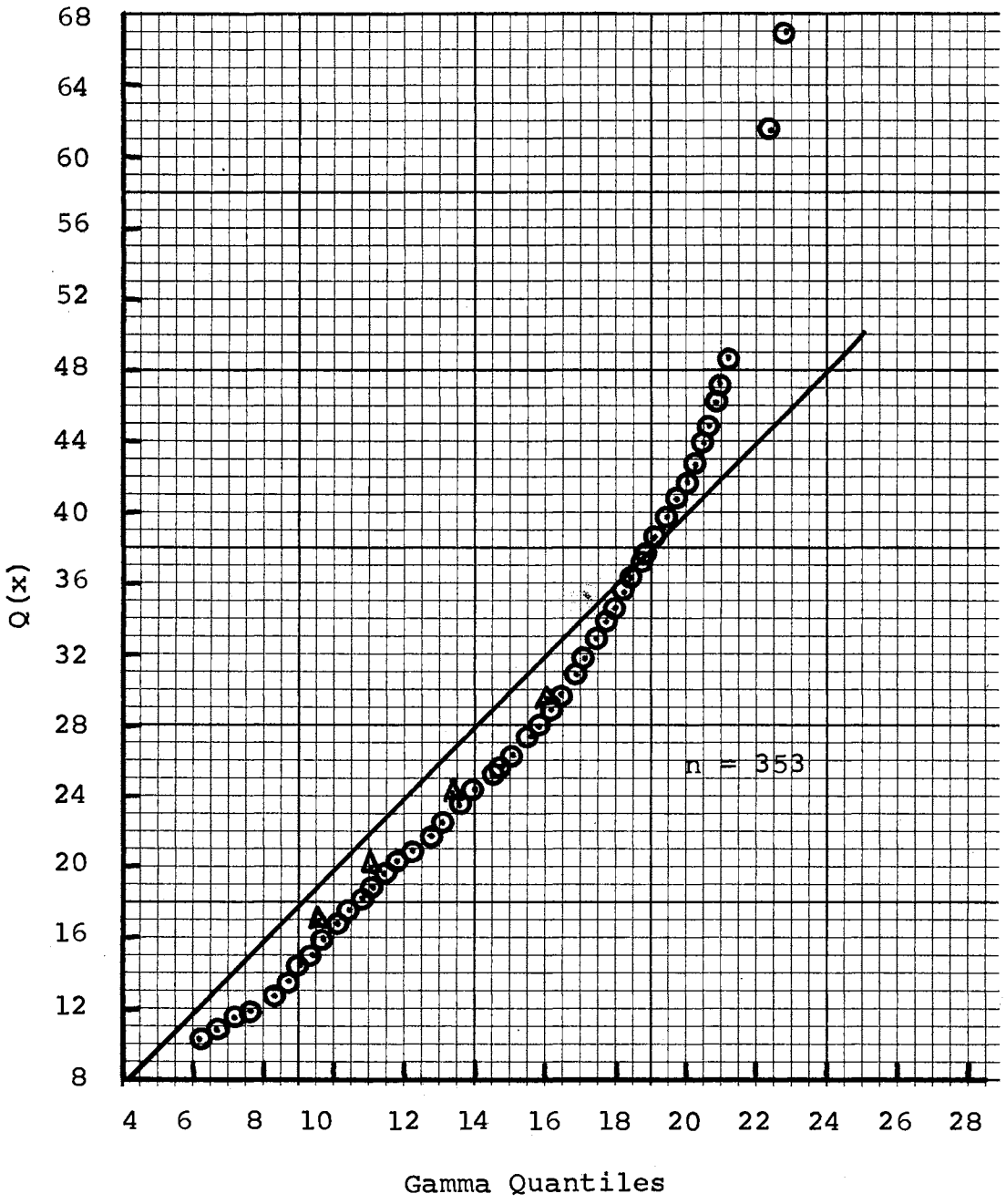


Figure 44. Multivariate Check for Age Group 45-49

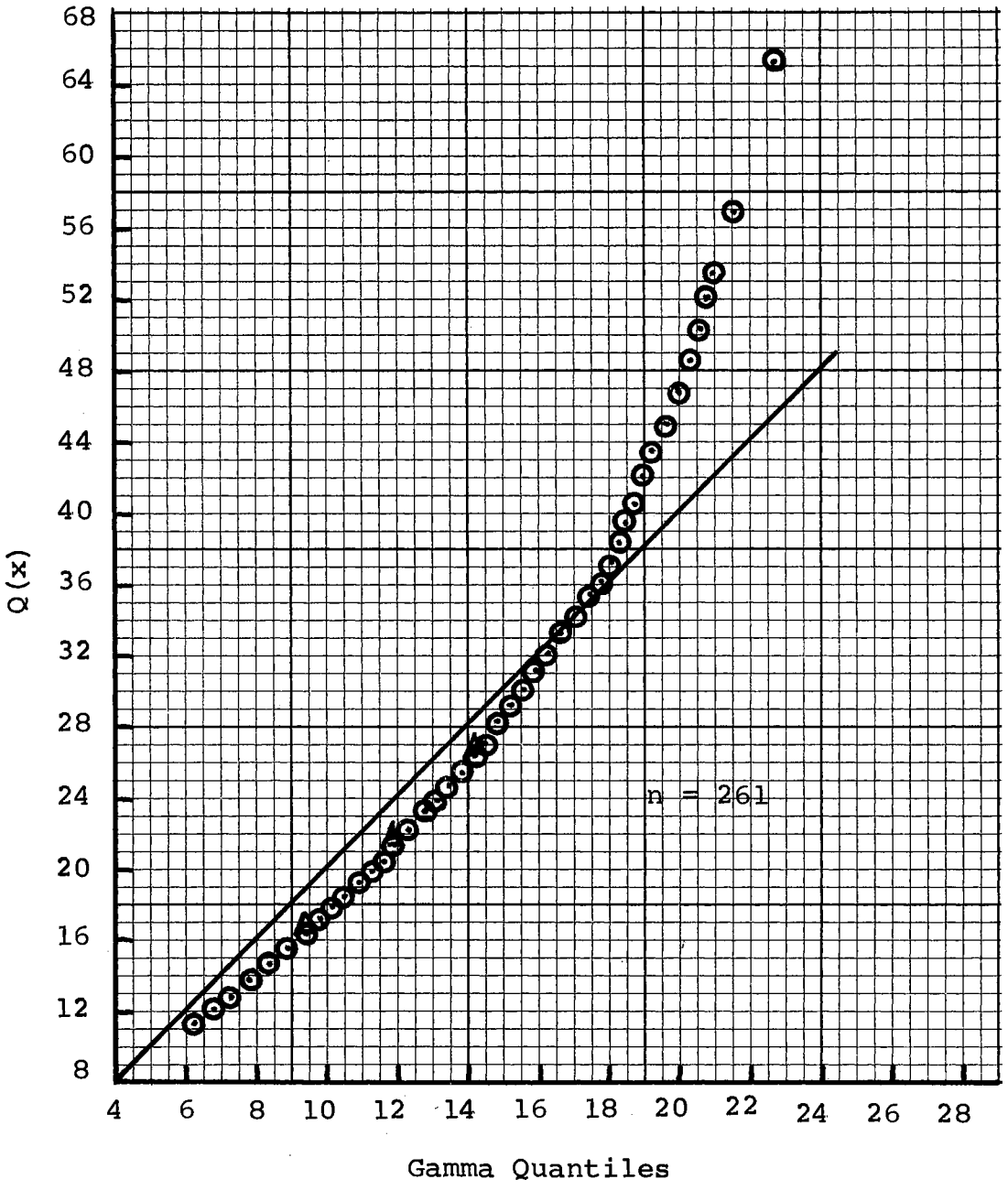


Figure 45. Multivariate Check for Age Group 50-54

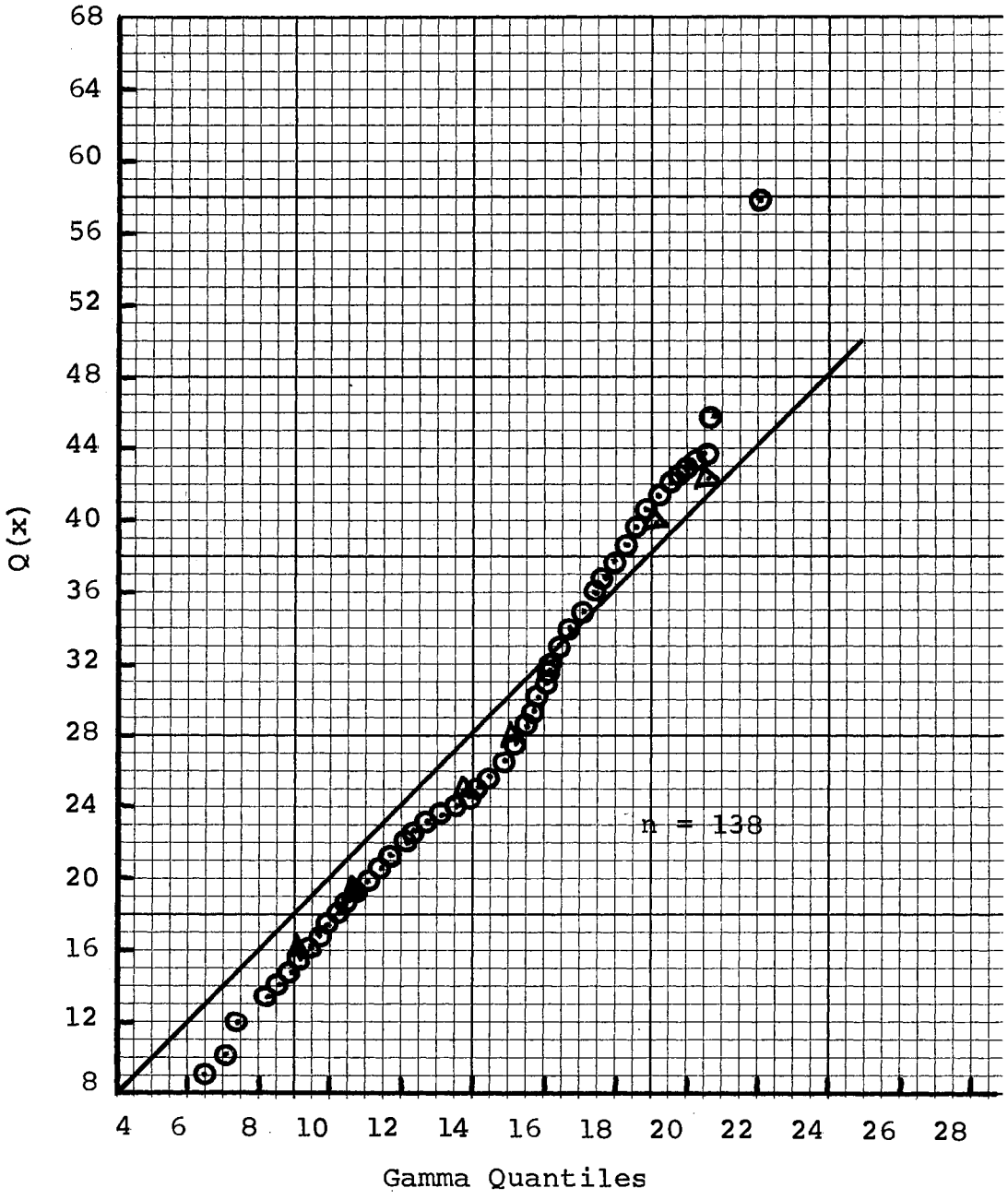


Figure 46. Multivariate Check for Age Group 55-59

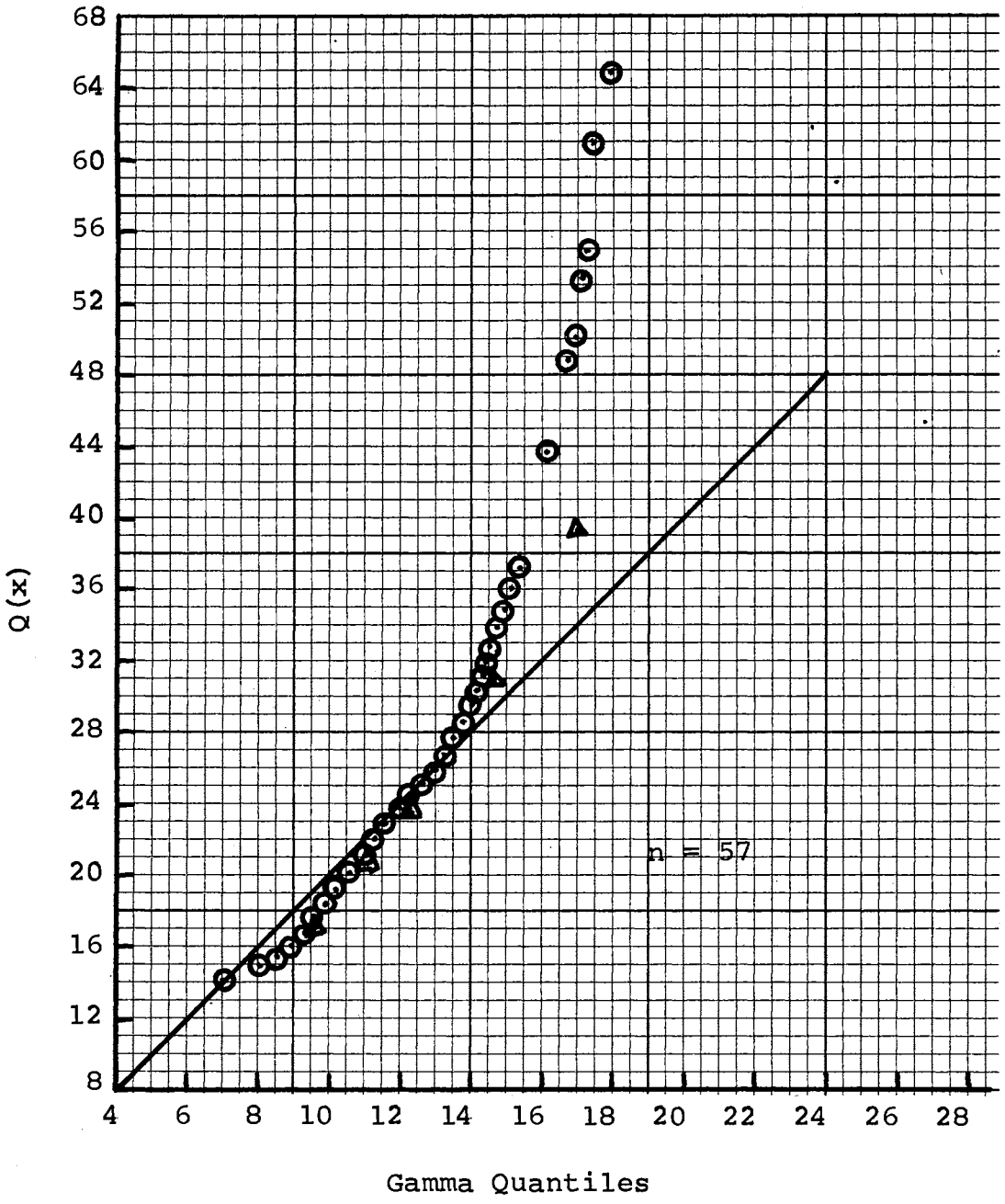


Figure 47. Multivariate Check for Age Group 60-64

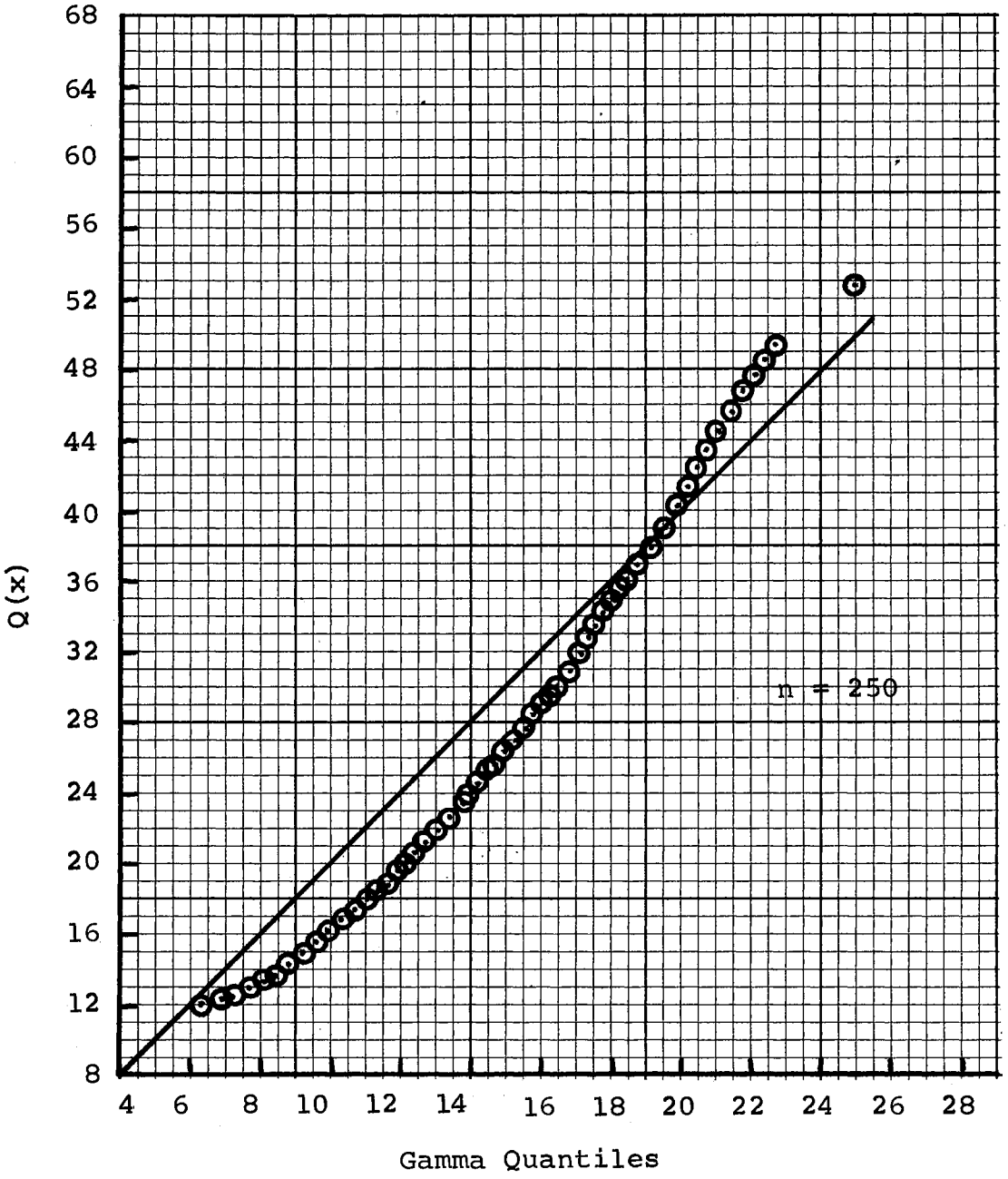


Figure 48. Multivariate Check for Age Group 50-54 Using Covariance for Age Group 50-64

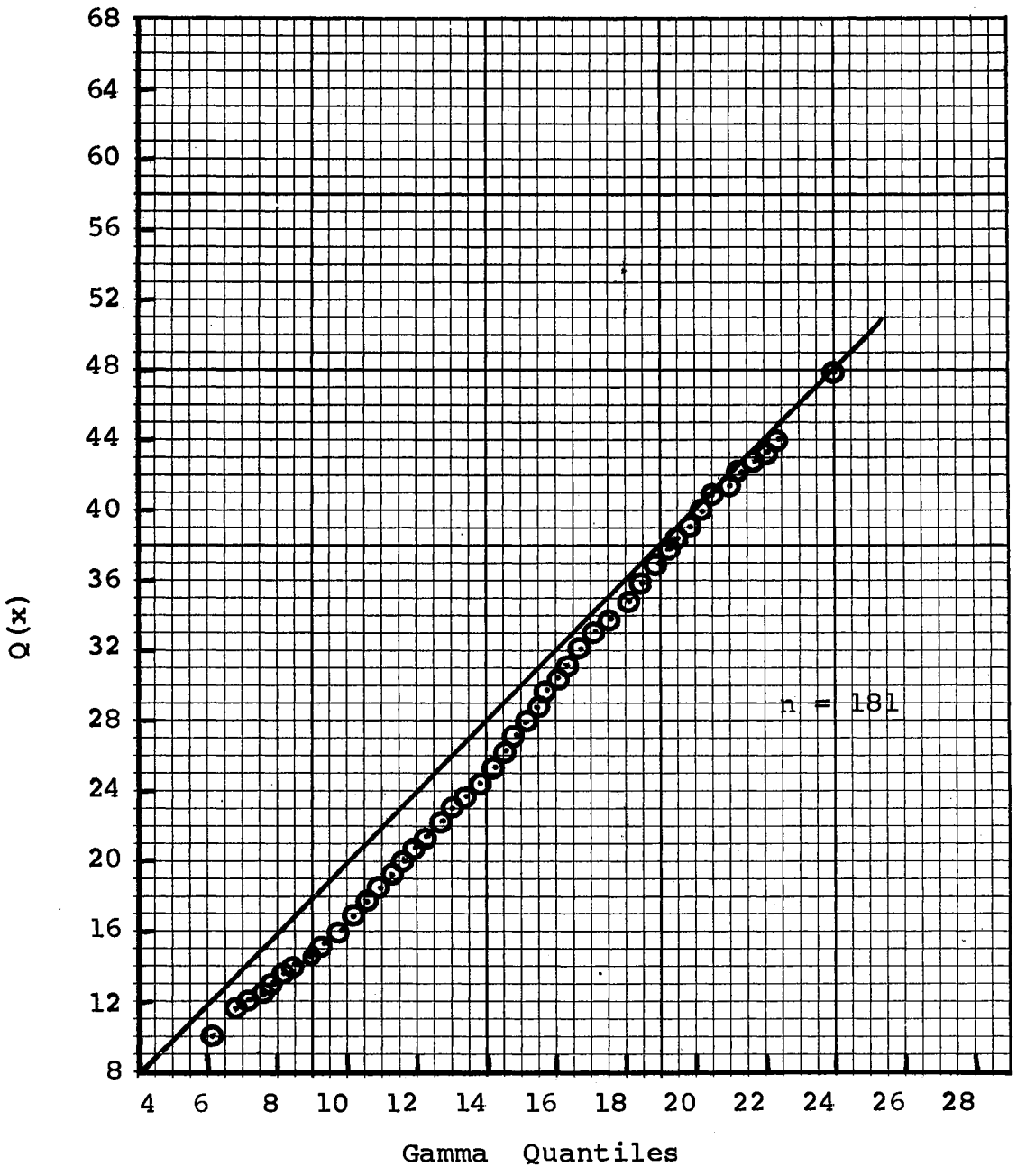


Figure 49. Multivariate Check for Age Group 55-64 Using Covariance for Age Group 50-64

APPENDIX E

AGE GROUP COVARIANCE MATRICES

TABLE XLVI

COVARIANCE MATRIX FOR AGE GROUP 20-29

ROW NO. 1									
6.2623	29.6419	.0180	-.0027	.0788	.0540	.1809	.2250	.5215	.0223
-.0145	-.0338	.0565	.1405	-.1726	.0000	-.5540	.2655	.0195	1.0481
.8952	.0133	-.0096	.0049	-.0054	2.3262	-.0781	.1494		
ROW NO. 2									
29.6419	583.7529	2.6293	-.3403	.8259	.5616	5.2384	5.9064	10.4841	.6718
.1104	.6068	.8517	3.3047	-2.7869	.3341	-6.8274	.8908	-.0322	20.1142
8.7967	.3348	-.4720	.1313	-.3083	12.6639	.0326	3.2350		
ROW NO. 3									
.0180	2.6293	.0150	-.0016	.0034	.0024	.0314	.0353	.0440	.0039
.0014	.0054	.0041	.0208	-.0190	.0000	-.0283	-.0011	-.0005	.0990
.0437	.0017	-.0018	.0005	-.0020	-.1165	.0012	.0131		
ROW NO. 4									
-.0027	-.3403	-.0016	.0035	.0020	.0025	.0186	.0102	.0129	.0020
.0005	.0007	.0018	.0040	.0057	.0000	-.0141	.0090	.0021	-.0201
.0093	-.0004	.0098	.0004	.0000	.2430	-.0031	-.0000		
ROW NO. 5									
.0788	.8259	.0034	.0020	.0135	.0105	.0086	.0170	.0344	.0036
-.0006	-.0025	-.0001	.0044	.0047	-.0015	-.0196	.0059	-.0002	.0573
-.0400	.0020	.0050	-.0020	-.0000	.2610	-.0063	-.0020		
ROW NO. 6									
.0540	.5616	.0024	.0025	.0105	.0175	.0046	.0139	.0604	.0047
-.0008	.0012	-.0006	.0119	.0335	.0058	.0278	.0071	.0000	.0407
-.0454	.0017	.0033	-.0029	-.0001	-.5322	-.0035	-.0054		
ROW NO. 7									
.1809	5.2384	.0314	.0186	.0086	.0046	.2492	.1407	.0571	.0002
.0116	.0142	.0299	.0616	-.1127	-.0294	-.1413	.0246	.0162	.1970
.2681	.0028	.0018	-.0029	-.0020	-2.0121	.0278	.0593		
ROW NO. 8									
.2250	5.9064	.0353	.0102	.0170	.0139	.1407	.1946	.1158	.0179
.0020	.0164	.0078	.0290	-.1610	-.0266	-.1099	.0376	.0084	.2666
.0849	.0068	.0043	-.0099	-.0048	-.0697	-.0257	.0107		
ROW NO. 9									
.5215	10.4841	.0440	.0329	.0344	.0604	.0571	.1158	4.1917	.0912
.0097	.0593	-.0291	.0192	-.5855	.0525	.3440	.2326	-.0297	-.0301
-.5725	.0043	.0276	.0043	-.0061	.3767	-.0165	-.1186		
ROW NO. 10									
.0223	.6718	.0039	.0020	.0036	.0047	.0002	.0179	.0912	.0319
.0023	.0034	.0003	.0159	-.0599	.0006	.0126	.0248	.0001	-.0102
-.0367	.0001	.0048	.0019	-.0010	-.1817	.0043	.0104		
ROW NO. 11									

TABLE XLVI (continued)

-.0145	.1104	.0014	.0005	-.0006	-.0008	.0116	.0020	.0097	.0023
.0139	.0087	.0027	.0031	-.0282	-.0083	-.0093	-.0137	.0042	-.0074
.0234	-.0006	.0019	.0008	-.0008	-.0071	.0057	.0044		
RCW NO. 12									
-.0338	.6068	.0054	.0007	-.0025	.0012	.0142	.0164	.0593	.0034
.0087	.0350	.0045	.0155	-.0753	.0006	.0144	-.0027	.0010	.0491
.0170	.0016	.0026	-.0013	.0008	-.9646	-.0223	-.0195		
RCW NO. 13									
.0565	.8517	.0041	.0018	-.0001	-.0006	.0299	.0078	-.0291	.0003
.0027	.0045	.0033	.0216	-.0250	-.0061	-.0226	.0122	.0013	.0165
.0414	-.0000	.0025	.0000	-.0002	-.4576	.0161	.0108		
RCW NO. 14									
.1405	3.3047	.0208	.0040	.0044	.0119	.0616	.0290	.0192	.0159
.0031	.0155	.0216	.3051	-.0344	-.0061	-.0172	-.0208	.0040	.1487
.1283	.0023	.0118	.0079	-.0010	-3.9620	-.0116	.0126		
RCW NO. 15									
-.1726	-2.7869	-.0190	.0057	.0447	.0335	-.1127	-.1610	-.5855	-.0599
-.0282	-.0753	-.0250	-.0344	3.0236	-.0346	.5163	-.1587	-.0058	-.3257
-.1894	-.0076	-.0238	-.0046	.0103	9.8198	.2049	.1168		
RCW NO. 16									
.0000	.3341	.0000	.0000	-.0015	.0058	-.0294	-.0266	.0525	.0006
-.0083	.0006	-.0061	-.0061	-.0346	.0838	.0718	.0118	.0008	-.0862
-.0943	-.0018	-.0004	-.0050	.0009	-1.1502	.0000	.0000		
RCW NO. 17									
-.5540	-6.8274	-.0283	-.0141	-.0196	.0278	-.1413	-.1099	.3440	.0126
-.0093	.0144	-.0226	-.0172	.5163	.0718	1.0699	.0400	.0046	-.2060
-.2400	-.0008	.0019	.0174	.0019	.7787	.0119	-.0988		
RCW NO. 18									
.2655	.8908	-.0011	.0090	.0059	.0071	.0246	.0376	.2326	.0248
-.0137	-.0027	.0122	-.0208	-.1587	.0118	.0400	.3072	-.0035	.0345
-.0519	.0048	.0092	-.0030	-.0022	-1.9839	-.0360	-.0417		
RCW NO. 19									
.0195	-.0322	-.0005	.0021	-.0002	.0000	.0162	.0084	-.0297	.0001
.0042	.0010	.0013	.0040	-.0058	.0008	.0096	-.0035	.0202	.0315
.0030	.0012	.0013	-.0013	.0009	.0324	-.0052	-.0095		
RCW NO. 20									
1.0481	20.1142	.0990	-.0201	.0573	.0407	.1970	.2666	-.0301	-.0102
-.0074	.0491	.0165	.1487	-.3257	-.0862	-.2060	.0345	.0315	2.1667
.7301	.0533	.0156	.0063	-.0065	-3.1884	-.2106	-.0238		
RCW NO. 21									
.8952	8.7967	.0437	.0093	-.0400	-.0454	.2681	.0849	-.5725	-.0367
.0234	.0170	.0414	.1283	-.1854	-.0943	-.2400	-.0519	.0030	.7301
2.3039	-.0081	-.0385	.0020	-.0173	-3.9546	-.0378	.0658		
RCW NO. 22									
.0133	.3348	.0017	-.0004	.0020	.0017	.0028	.0068	.0043	.0001

TABLE XLVI (continued)

	-.0006	.0016	-.0000	.0023	-.0076	-.0018	-.0008	.0048	.0012	.0533
	-.0081	.0019	.0010	.0002	.0003	-.2346	-.0080	-.0063		
RCW NO. 23										
	-.0096	-.4720	-.0018	.0098	.0050	.0033	.0018	.0043	.0276	.0048
	.0019	.0026	.0025	.0118	-.0238	-.0004	.0019	.0092	.0013	.0156
	-.0385	.0010	.0203	.0001	.0014	-.2839	-.0166	-.0191		
RCW NO. 24										
	.0049	.1313	.0005	.0004	-.0020	-.0029	-.0029	-.0099	.0043	.0019
	.0008	-.0013	.0000	.0079	-.0146	-.0050	.0174	-.0030	-.0013	.0063
	.0020	.0002	.0001	.0141	-.0005	.5806	.0083	.0166		
RCW NO. 25										
	-.0054	-.3083	-.0020	.0000	-.0000	-.0001	-.0020	-.0048	-.0061	-.0010
	-.0008	.0008	-.0002	-.0010	.0103	.0009	.0019	-.0022	.0009	-.0065
	-.0173	.0003	.0014	-.0005	.0025	-.1613	-.0002	-.0014		
RCW NO. 26										
	2.3262	12.6639	-.1165	.2430	.2610	-.5322	-2.0121	-.0697	.3767	-.1817
	-.0071	-.9646	-.4576	-3.9620	9.8198	-1.1502	.7787	-1.9839	.0324	-3.1884
	-3.9546	-.2346	-.2839	.5806	-.1613	712.8091	.2426	1.6377		
RCW NO. 27										
	-.0781	.0326	.0012	-.0031	-.0063	-.0035	.0278	-.0257	-.0165	.0043
	.0057	-.0223	.0161	-.0116	.2049	.0000	.0119	-.0360	-.0052	-.2106
	-.0378	-.0080	-.0166	.0083	-.0002	.2426	.1140	.0943		
RCW NO. 28										
	.1494	3.2350	.0131	-.0000	-.0020	-.0054	.0593	.0107	-.1186	.0104
	.0044	-.0195	.0108	.0126	.1168	.0000	-.0988	-.0417	-.0095	-.0238
	.0658	-.0063	-.0191	.0066	-.0014	1.6377	.0943	.1493		

TABLE XLVII

COVARIANCE MATRIX FOR AGE GROUP 30-34

ROW NO. 1									
5.4296	23.7484	-.0192	-.0272	.0157	.0107	.0113	-.0049	.6412	-.0489
-.0391	-.0517	.0166	-.0266	-.1558	-.0152	-.0751	-.0498	-.0135	.3544
1.2796	-.0044	-.0300	.0199	.0126	-3.3156	1.0208	.9725		
ROW NO. 2									
23.7484	511.2704	2.1614	.0831	.8289	.9301	3.1539	4.7728	7.2009	.1051
-.2997	.4803	.7132	2.3697	-.1474	-.0633	-2.0243	-.0956	-.0878	12.6001
17.5882	.2058	.0617	.6137	.0760	-197.5543	7.1293	5.9174		
ROW NO. 3									
-.0192	2.1614	.0125	.0010	.0040	.0049	.0182	.0276	.0226	.0015
-.0005	.0039	.0029	.0139	.0014	.0001	-.0098	.0006	-.0002	.0617
.0650	.0013	.0008	.0031	.0001	-1.0119	.0109	.0056		
ROW NO. 4									
-.0272	.0831	.0010	.0156	.0024	.0027	-.0017	.0046	.0328	.0069
.0015	.0016	-.0006	.0120	.0139	-.0019	.0394	-.0073	-.0016	-.0209
-.0113	-.0005	.0145	.0014	-.0010	-.0237	-.0480	-.0368		
ROW NO. 5									
.0157	.8289	.0040	.0024	.0131	.0111	.0003	.0096	.0363	.0031
.0001	.0026	-.0015	.0110	.0234	.0042	.0173	-.0072	-.0002	.0190
.0190	.0005	.0050	.0005	-.0009	-.3922	-.0013	.0052		
ROW NO. 6									
.0107	.9301	.0049	.0027	.0111	.0176	.0117	.0187	.0554	.0038
-.0008	.0037	-.0003	.0094	.0189	.0012	.0105	-.0005	-.0017	.0302
-.0101	.0012	.0042	.0002	-.0012	-.6165	.0151	.0090		
ROW NO. 7									
.0113	3.1539	.0182	-.0017	.0003	.0117	.2365	.1255	.1744	-.0021
.0034	.0056	.0010	-.0029	-.0599	.0080	-.0409	.0248	.0047	.2160
.0884	.0052	.0069	.0168	.0026	-1.3651	.0459	.0401		
ROW NO. 8									
-.0049	4.7728	.0276	.0046	.0096	.0187	.1255	.1523	.0319	-.0028
-.0008	-.0010	.0046	.0483	.1062	-.0036	.0067	.0153	-.0057	.1733
.0895	.0044	.0054	.0120	.0012	-1.2087	.0262	.0272		
ROW NO. 9									
.6412	7.2009	.0226	.0328	.0363	.0554	.1744	.0319	5.3676	.1674
.0322	.0578	.0572	.0657	-1.2306	.0949	-.2056	-.0890	.0226	.5321
.0422	.0171	.0504	-.0035	-.0087	-7.8829	-.5978	-.3642		
ROW NO. 10									
-.0489	.1051	.0015	.0069	.0031	.0038	-.0021	-.0028	.1674	.0401
.0004	.0046	-.0008	-.0016	-.0783	.0021	-.0032	-.0058	.0013	-.0057
-.0187	.0002	.0097	.0016	-.0004	-.2280	-.0239	-.0193		
ROW NO. 11									

TABLE XLVII (continued)

-.0391	-.2997	-.0005	.0015	.0001	-.0008	.0034	-.0008	.0322	.0004
.0155	.0015	.0032	.0041	-.0188	-.0001	.0052	-.0030	.0018	-.0070
.0039	-.0003	.0028	.0007	-.0000	-.0444	-.0206	-.0104		
RCW NO. 12									
-.0517	.4803	.0039	.0016	.0026	.0037	.0056	-.0010	.0578	.0046
.0015	.0324	.0071	.0245	-.0586	.0037	.0348	.0114	.0028	.0138
-.0232	.0008	.0006	-.0025	-.0010	-.8973	-.0074	-.0226		
RCW NO. 13									
.0166	.7132	.0029	-.0006	-.0015	-.0003	.0010	.0046	.0572	-.0008
.0032	.0071	.0474	.0042	-.0208	-.0094	-.0341	.0125	-.0008	.0386
.0030	.0012	-.0013	.0005	-.0007	-1.0017	.0254	.0198		
RCW NO. 14									
-.0266	2.3697	.0134	.0120	.0110	.0094	-.0029	.0483	.0657	-.0016
.0041	.0245	.0042	.3422	.3125	-.0078	.2066	-.0227	-.0069	.0965
-.2137	.0062	.0161	.0014	-.0015	.2846	-.1661	-.1172		
RCW NO. 15									
-.1558	-.1474	.0014	.0139	.0234	.0189	-.0599	.1062	-1.2306	-.0783
-.0188	-.0586	-.0208	.3125	3.3791	-.0726	.3581	-.0728	-.0197	.3258
.0794	.0107	.0140	.0378	.0001	2.3661	.4858	.4828		
RCW NO. 16									
-.0152	-.0633	.0001	-.0019	.0042	.0012	.0080	-.0036	.0949	.0021
-.0001	.0037	-.0094	-.0078	-.0726	.0385	.0248	-.0098	.0078	-.0181
-.0047	-.0005	-.0006	-.0003	-.0001	-.3191	-.0802	-.0410		
RCW NO. 17									
-.0751	-2.0243	-.0098	.0394	.0173	.0105	-.0409	.0067	-.2056	-.0032
.0052	.0348	-.0341	.2066	.3581	.0248	2.6683	.0880	.0056	-.2283
-.2020	-.0039	.0227	-.0558	-.0077	2.5108	-.0820	-.0001		
RCW NO. 18									
-.0498	-.0956	.0006	-.0073	-.0072	-.0005	.0248	.0153	-.0890	-.0058
-.0030	.0114	.0125	-.0227	-.0728	-.0098	.0880	.3816	-.0199	.0629
-.0283	.0022	-.0092	.0001	.0001	-.6002	.1295	.0774		
RCW NO. 19									
-.0135	-.0878	-.0002	-.0016	-.0002	-.0017	.0047	-.0057	.0226	.0013
.0018	.0028	-.0008	-.0069	-.0197	.0078	.0056	-.0199	.0236	-.0048
.0462	-.0005	-.0014	.0008	-.0000	-.5234	.0156	-.0058		
RCW NO. 20									
.3544	12.6001	.0617	-.0204	.0190	.0302	.2160	.1733	.5321	-.0057
-.0070	.0138	.0386	.0865	.3258	-.0181	-.2283	.0629	-.0048	1.1837
.8018	.0278	-.0105	.0264	.0034	-11.4774	.3236	.1879		
RCW NO. 21									
1.2796	17.5882	.0650	-.0113	.0190	-.0101	.0884	.0895	.0422	-.0187
.0039	-.0232	.0030	-.2137	.0794	-.0047	-.2020	-.0283	.0462	.8018
2.1256	-.0006	-.0182	.0403	.0003	-13.0014	.6833	.5451		
RCW NO. 22									
-.0044	.2058	.0013	-.0005	.0005	.0012	.0052	.0044	.0171	.0002

TABLE XLVII (continued)

-.0003	.0008	.0012	.0062	.0107	-.0005	-.0039	.0022	-.0005	.0278
-.0006	.0009	-.0002	.0003	.0001	-.2243	.0036	.0008		
RCN NO. 23									
-.0300	.0617	.0008	.0145	.0050	.0042	.0064	.0054	.0504	.0097
.0028	.0006	-.0013	.0161	.0140	-.0006	.0227	-.0092	-.0014	-.0105
-.0182	-.0002	.0271	.0012	-.0021	-.5084	-.0330	-.0266		
RCN NO. 24									
.0199	.6137	.0031	.0014	.0005	.0002	.0168	.0120	-.0035	.0016
.0007	-.0025	.0005	.0014	.0378	-.0003	-.0558	.0001	.0008	.0264
.0403	.0003	.0012	.0151	.0004	-.5014	.0224	.0160		
RCN NO. 25									
.0126	.0760	.0001	-.0010	-.0009	-.0012	.0026	.0012	-.0087	-.0004
-.0000	-.0010	-.0007	-.0015	.0001	-.0001	-.0077	.0001	-.0000	.0034
.0003	.0001	-.0021	.0004	.0020	.2421	-.0002	.0039		
RCN NO. 26									
-3.3156	-197.5543	-1.0119	-.0237	-.3922	-.6165	-1.3651	-1.2087	-7.8829	-.2280
-.0444	-.8973	-1.0017	.2846	2.3661	-.3191	2.5108	-.6002	-.5234	-11.4774
-13.0014	-.2243	-.5084	-.5014	.2421	859.3680	-3.0118	-2.9971		
RCN NO. 27									
1.0208	7.1293	.0109	-.0480	-.0013	.0151	.0459	.0262	-.5978	-.0239
-.0206	-.0074	.0254	-.1661	.4858	-.0802	-.0820	.1295	.0156	.3236
.6833	.0036	-.0330	.0224	-.0002	-3.0118	.5265	.3954		
RCN NO. 28									
.9725	5.9174	.0056	-.0368	.0052	.0090	.0401	.0272	-.3642	-.0193
-.0104	-.0226	.0198	-.1172	.4828	-.0410	-.0001	.0774	.0058	.1879
.5451	.0008	-.0266	.0160	.0039	-2.9971	.3954	.3649		

TABLE XLVIII

COVARIANCE MATRIX FOR AGE GROUP 35-39

ROW NO. 1									
5.6548	28.1657	-.0022	-.0133	.0245	.0207	.1659	.0301	-.2379	-.0698
.0013	-.0323	.0275	-.0104	.1514	.0177	.1375	.0424	-.0109	.1151
1.4745	-.0165	-.0099	.0805	.0055	6.0476	.9019	.6107		
ROW NO. 2									
28.1657	570.5414	2.3167	.0943	.8048	1.0706	4.1047	4.9639	1.6623	.0666
.3136	.2263	.8710	2.5428	2.0794	.5910	-.2387	1.6146	-.3403	19.8084
25.5868	.3053	.3381	.3569	-.1394	-187.9740	5.6166	2.9480		
ROW NO. 3									
-.0022	2.3167	.0128	.0007	.0039	.0053	.0187	.0283	.0167	.0025
.0017	.0024	.0043	.0140	.0066	.0021	-.0092	.0070	-.0015	.1085
.0962	.0022	.0020	.0005	-.0010	-1.2628	.0050	-.0020		
ROW NO. 4									
-.0133	.0943	.0007	.0162	.0023	.0029	-.0009	.0022	.0565	.0041
.0017	.0021	.0012	.0102	.0032	.0020	.0050	.0000	-.0005	-.0122
-.0159	-.0002	.0139	-.0027	.0001	.1360	-.0103	-.0108		
ROW NO. 5									
.0245	.8048	.0039	.0023	.0121	.0093	.0083	.0096	.0229	.0020
.0016	.0013	.0020	.0118	.0401	.0015	.0130	.0014	-.0003	.0402
.0337	.0009	.0053	.0011	-.0001	-.5238	.0067	.0055		
ROW NO. 6									
.0207	1.0706	.0053	.0029	.0093	.0163	.0129	.0159	.0334	.0006
.0013	.0021	.0010	.0119	.0289	.0007	.0115	.0025	-.0007	.0556
.0410	.0013	.0049	.0026	-.0004	-.6045	.0027	-.0018		
ROW NO. 7									
.1659	4.1047	.0187	-.0009	.0083	.0129	.2219	.1040	.0300	.0076
.0086	.0094	.0296	.0294	.0151	-.0048	-.0418	.0369	-.0116	.0923
.1205	.0012	.0061	.0021	-.0014	-.9428	.0599	.0361		
ROW NO. 8									
.0301	4.9639	.0283	.0022	.0096	.0159	.1040	.1510	.0311	.0105
.0065	.0110	.0204	.0118	.0750	.0030	-.0450	.0072	-.0071	.1942
.1636	.0042	.0052	.0037	-.0036	-3.1475	.0036	-.0170		
ROW NO. 9									
-.2379	1.6623	.0167	.0565	.0229	.0334	.0300	.0311	5.1464	.0856
.0120	.0343	-.0023	.1347	-.2457	-.0083	-.0472	-.0072	.0043	.2205
.0719	.0070	.0557	-.0280	.0011	-4.6116	-.2718	-.1852		
ROW NO. 10									
-.0698	.0666	.0025	.0041	.0020	.0006	.0076	.0105	.0856	.0482
.0033	.0035	.0038	.0120	-.0075	-.0027	-.0123	-.0042	-.0002	.0317
.0049	.0009	.0069	-.0047	-.0001	-.7481	.0002	-.0023		
ROW NO. 11									

.0013	.3136	.0017	.0017	.0016	.0013	.0086	.0065	.0120	.0033
.0202	.0041	.0065	.0098	-.0152	-.0011	-.0293	-.0018	.0002	.0067
.0095	.0001	.0010	-.0003	.0001	-.2074	-.0054	.0013		
RCW NO. 12									
-.0323	.2263	.0024	.0021	.0013	.0021	.0094	.0110	.0343	.0035
.0041	.0275	.0033	.0182	-.0308	-.0002	-.0110	.0031	-.0007	.0217
.0188	.0005	.0033	-.0021	-.0006	-.4060	-.0175	-.0165		
RCW NO. 13									
.0275	.8710	.0043	.0012	.0020	.0010	.0296	.0204	-.0023	.0038
.0065	.0033	.0532	.0167	.0304	-.0071	-.0439	.0048	-.0028	-.0078
.0295	-.0007	.0024	-.0012	-.0005	.0218	.0358	.0214		
RCW NO. 14									
-.0104	2.5428	.0140	.0102	.0118	.0119	.0294	.0118	.1347	.0120
.0098	.0182	.0167	.2883	.1407	.0022	-.0135	.0130	-.0099	.0772
.0820	.0014	.0209	-.0082	.0007	-1.2191	-.0426	-.0549		
RCW NO. 15									
.1514	2.0794	.0066	.0032	.0401	.0289	.0191	.0750	-.2457	-.0075
-.0152	-.0308	.0304	.1407	2.9559	-.0072	.7439	-.0721	-.0040	-.1128
.1300	-.0051	.0091	.0062	-.0008	-3.0850	.3126	.1585		
RCW NO. 16									
.0177	.5910	.0021	.0020	.0015	.0007	-.0048	.0030	-.0083	-.0027
-.0011	-.0002	-.0071	.0022	-.0072	.0806	.0490	.0015	.0057	.0213
.0181	.0004	.0013	.0042	-.0007	.0984	.0000	.0000		
RCW NO. 17									
.1375	-.2387	-.0092	.0050	.0130	.0115	-.0418	-.0450	-.0472	-.0123
-.0293	-.0110	-.0439	-.0135	.7439	.0490	1.5704	.0617	-.0113	-.1196
-.0993	-.0021	.0035	-.0141	.0010	.8826	.0220	.0479		
RCW NO. 18									
.0424	1.6146	.0070	.0000	.0014	.0025	.0369	.0072	-.0072	-.0042
-.0018	.0031	.0048	.0130	-.0721	.0015	.0617	.3536	-.0143	-.0204
.0066	-.0004	.0014	-.0072	-.0024	-.0894	-.0317	-.0583		
RCW NO. 19									
-.0109	-.3403	-.0015	-.0005	-.0003	-.0007	-.0116	-.0071	.0043	-.0002
.0002	-.0007	-.0028	-.0099	-.0040	.0057	-.0113	-.0143	.0239	-.0084
-.0216	.0000	-.0007	-.0006	.0002	.0313	-.0049	-.0088		
RCW NO. 20									
.1151	19.8084	.1085	-.0122	.0402	.0556	.0923	.1942	.2205	.0317
.0067	.0217	-.0078	.0772	-.1128	.0213	-.1196	-.0209	-.0084	1.9037
1.2566	.0449	-.0184	.0281	.0026	-17.0913	-.1099	-.1575		
RCW NO. 21									
1.4745	25.5868	.0962	-.0159	.0337	.0410	.1205	.1636	.0719	.0049
.0095	.0188	.0295	.0820	.1300	.0181	-.0993	.0066	-.0216	1.2566
3.1016	.0001	-.0160	.0301	-.0006	-9.5117	.4517	.1898		
RCW NO. 22									
-.0165	.3053	.0022	-.0002	.0009	.0013	.0012	.0042	.0070	.0009

TABLE XLVIII (continued)

.0001	.0005	-.0007	.0014	-.0051	.0004	-.0021	-.0004	.0000	.0449
.0001	.0014	-.0003	.0005	.0001	-.4324	-.0093	-.0076		
RCW NO. 23									
-.0099	.3381	.0020	.0139	.0053	.0049	.0061	.0052	.0557	.0069
.0010	.0033	.0024	.0209	.0091	.0013	.0035	.0014	-.0007	-.0184
-.0160	-.0003	.0257	-.0031	-.0005	.0600	-.0151	-.0152		
RCW NO. 24									
.0605	.3969	.0005	-.0027	.0011	.0026	.0021	.0037	-.0280	-.0047
-.0003	-.0021	-.0012	-.0082	.0062	.0042	-.0141	-.0072	-.0006	.0281
.0301	.0005	-.0031	.0153	-.0001	-.3850	.0192	.0113		
RCW NO. 25									
.0055	-.1394	-.0010	.0001	-.0001	-.0004	-.0014	-.0036	.0011	-.0001
.0001	-.0006	-.0005	.0007	-.0008	-.0007	.0010	-.0024	.0002	.0026
-.0006	.0001	-.0005	-.0001	.0023	.0203	-.0083	-.0048		
RCW NO. 26									
6.0476	-187.9740	-1.2628	.1360	-.5238	-.6045	-.9428	-3.1475	-4.6116	-.7481
-.2074	-.4060	.0218	-1.2191	-3.0850	.0984	.8826	-.0894	.0313	-17.0913
-9.5117	-.4324	.0600	-.3850	.0203	899.8584	1.8055	2.8053		
RCW NO. 27									
.9019	5.6166	.0050	-.0103	.0067	.0027	.0599	.0036	-.2718	.0002
-.0054	-.0175	.0358	-.0426	.3126	.0000	.0220	-.0317	-.0049	-.1099
.4517	-.0093	-.0151	.0192	-.0083	1.8055	.4889	.3167		
RCW NO. 28									
.6107	2.9480	-.0020	-.0108	.0055	-.0018	.0361	-.0170	-.1852	-.0023
.0013	-.0165	.0214	-.0549	.1585	.0000	.0479	-.0583	-.0088	-.1575
.1898	-.0076	-.0152	.0913	-.0048	2.8053	.3167	.3090		

TABLE XLIX

COVARIANCE MATRIX FOR AGE GROUP 40-44

ROW NO. 1									
5.9899	28.3674	-.0047	-.0125	.0002	.0076	-.0007	.0819	-.2231	-.0019
-.0110	-.0103	.0189	.0338	.1463	.0000	.0219	.0272	.0056	.5469
1.2025	.0027	-.0231	.0013	-.0051	2.3443	.7244	.4498		
ROW NO. 2									
28.3674	554.8030	2.2265	.0145	.6480	.6198	3.7634	4.7937	-1.6415	-.0047
.3151	.1556	.6284	3.0431	2.9082	.0000	1.7685	.4893	-.0171	17.8004
22.5003	.2491	-.0013	.2681	-.0272	-128.3488	4.3519	2.7648		
ROW NO. 3									
-.0047	2.2265	.0125	.0003	.0033	.0031	.0208	.0243	-.0086	-.0001
.0021	.0010	.0031	.0162	.0147	.0000	.0085	.0028	-.0000	.0793
.0846	.0012	.0005	.0013	-.0002	-.8637	.0055	.0036		
ROW NO. 4									
-.0125	.0145	.0003	.0166	.0019	.0018	.0055	-.0015	.0519	.0042
.0010	.0005	.0039	.0107	.0206	.0000	-.0037	.0025	-.0016	-.0142
.0101	-.0005	.0153	-.0009	-.0000	.3495	.0145	.0041		
ROW NO. 5									
.0002	.6480	.0033	.0019	.0148	.0113	.0070	.0071	.0278	.0018
.0012	.0007	.0011	.0108	.0549	.0000	.0239	-.0023	.0016	.0463
.0269	.0010	.0049	-.0006	.0001	-.1894	-.0059	-.0083		
ROW NO. 6									
.0076	.6198	.0031	.0018	.0113	.0173	.0053	.0074	.0455	.0013
.0011	.0013	.0017	.0093	.0464	.0000	.0224	-.0009	.0013	.0434
.0132	.0011	.0031	-.0011	.0003	-.4733	-.0042	-.0062		
ROW NO. 7									
-.0007	3.7634	.0208	.0055	.0070	.0053	.2204	.0840	-.0336	-.0075
-.0013	.0018	.0210	.0367	-.0535	.0000	-.0001	.0187	-.0089	.0860
.0470	.0017	.0120	.0017	-.0030	-.9906	-.0123	-.0079		
ROW NO. 8									
.0819	4.7937	.0243	-.0015	.0071	.0074	.0840	.1200	.0016	-.0017
.0015	-.0008	.0086	.0292	-.0061	.0000	.0170	.0087	.0004	.1672
.0975	.0036	.0046	.0004	-.0016	-2.4880	-.0053	.0019		
ROW NO. 9									
-.2231	-1.6415	-.0086	.0519	.0278	.0455	-.0336	.0016	5.7430	.0835
.0263	.0654	.0057	.0900	-.4188	.0000	-.0793	.0696	-.0006	-.4042
-.6112	-.0061	.0796	-.0527	.0108	-2.3308	-.0917	-.2215		
ROW NO. 10									
-.0019	-.0047	-.0001	.0042	.0018	.0013	-.0075	-.0017	.0835	.0431
.0016	.0030	-.0004	.0058	-.0182	.0000	-.0130	.0021	.0025	-.0054
.0110	-.0004	.0072	-.0018	-.0012	-.3915	-.0002	-.0150		
ROW NO. 11									

TABLE XLIX (continued)

-.0110	.3151	.0021	.0010	.0012	.0011	-.0013	.0015	.0263	.0016
.0241	.0031	.0071	.0160	-.0151	.0000	-.0030	-.0005	.0016	.0197
.0219	.0003	.0032	-.0000	.0004	.2754	-.0052	-.0017		
ROW NO. 12									
-.0103	.1556	.0010	.0005	.0007	.0013	.0018	-.0008	.0654	.0030
.0031	.0303	.0069	.0196	-.0388	.0000	-.0138	.0032	-.0013	-.0035
.0039	-.0001	.0013	-.0011	.0010	-.2303	-.0129	-.0148		
ROW NO. 13									
.0189	.6284	.0031	.0039	.0011	.0017	.0210	.0086	.0057	-.0004
.0071	.0069	.0508	.0168	.0154	.0000	-.0054	.0068	-.0016	-.0003
.0596	-.0008	.0058	-.0001	-.0001	-.0994	.0199	.0061		
ROW NO. 14									
.0338	3.0431	.0162	.0107	.0108	.0093	.0367	.0292	.0900	.0058
.0160	.0196	.0168	.3431	.1677	.0000	.0650	.0202	-.0113	.0768
.0962	.0011	.0180	-.0023	.0020	-2.6551	-.0024	-.0153		
ROW NO. 15									
.1463	2.9082	.0147	.0206	.0549	.0464	-.0535	-.0061	-.4188	-.0182
-.0151	-.0388	.0154	.1677	3.1142	.0000	.8829	.0189	-.0232	-.0587
-.2345	.0010	.0240	-.0097	.0098	2.1306	.0326	-.0166		
ROW NO. 16									
.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		
ROW NO. 17									
.0219	1.7685	.0085	-.0037	.0239	.0224	-.0001	.0170	-.0793	-.0130
-.0030	-.0138	-.0054	.0650	.8829	.0000	1.3258	.1321	-.0324	.0247
-.1722	.0034	-.0085	.0005	-.0029	.0565	-.1016	-.0654		
ROW NO. 18									
.0272	.4893	.0028	.0025	-.0023	-.0009	.0187	.0087	.0696	.0021
-.0005	.0032	.0068	.0202	.0189	.0000	.1321	.3667	-.0074	.0237
.0771	-.0001	.0032	.0022	-.0015	.3541	-.0319	-.0430		
ROW NO. 19									
.0056	-.0171	-.0000	-.0016	.0016	.0013	-.0089	.0004	-.0006	.0025
.0016	-.0013	-.0016	-.0113	-.0232	.0000	-.0324	-.0074	.0250	.0217
.0261	.0003	-.0008	.0004	.0006	.4061	.0014	.0078		
ROW NO. 20									
.5469	17.8004	.0793	-.0142	.0463	.0434	.0860	.1672	-.4042	-.0054
.0197	-.0035	-.0003	.0768	-.0587	.0000	.0247	.0237	.0217	1.9860
1.2284	.0443	-.0238	.0047	.0005	-9.6844	-.0105	.0027		
ROW NO. 21									
1.2025	22.5003	.0846	.0101	.0269	.0132	.0470	.0975	-.6112	.0110
.0219	.0039	.0596	.0962	-.2345	.0000	-.1722	.0771	.0261	1.2284
3.3198	-.0052	.0087	.0264	.0001	1.6403	.4878	.3383		
ROW NO. 22									
.0027	.2491	.0012	-.0005	.0010	.0011	.0017	.0036	-.0061	-.0004

TABLE XLIX (continued)

.0003	-.0001	-.0008	.0011	.0010	.0000	.0034	-.0001	.0003	.0443
-.0052	.0015	-.0009	-.0001	.0000	-.3291	-.0069	-.0045		
RCW NO. 23									
-.0231	-.0013	.0005	.0153	.0049	.0031	.0120	.0046	.0796	.0072
.0032	.0013	.0058	.0180	.0240	.0000	-.0085	.0032	-.0008	-.0238
.0087	-.0009	.0277	-.0018	-.0001	.1735	.0146	-.0016		
RCW NO. 24									
.0013	.2681	.0013	-.0009	-.0006	-.0011	.0017	.0004	-.0527	-.0018
-.0000	-.0011	-.0001	-.0023	-.0097	.0000	.0005	.0022	.0004	.0047
.0264	-.0001	-.0018	.0135	-.0008	-.4215	.0082	.0066		
RCW NO. 25									
-.0051	-.0272	-.0002	-.0000	.0001	.0003	-.0030	-.0016	.0108	-.0012
.0004	.0010	-.0001	.0020	.0098	.0000	-.0029	-.0015	.0006	.0005
.0001	.0000	-.0001	-.0008	.0025	.0059	-.0014	-.0007		
RCW NO. 26									
2.3443	-128.3488	-.8637	.3495	-.1894	-.4733	-.9906	-2.4880	-2.3308	-.3915
.2754	-.2303	-.0994	-2.6551	2.1306	.0000	.0565	.3541	.4061	-9.6844
1.6403	-.3291	.1735	-.4215	.0059	926.4102	1.4077	-.3399		
RCW NO. 27									
.7244	4.3519	.0055	.0145	-.0059	-.0042	-.0123	-.0053	-.0917	-.0002
-.0052	-.0129	.0199	-.0024	.0326	.0000	-.1016	-.0319	.0014	-.0105
.4878	-.0069	.0146	.0082	-.0014	1.4077	.3891	.2329		
RCW NO. 28									
.4498	2.7648	.0036	.0041	-.0083	-.0062	-.0079	.0019	-.2215	-.0150
-.0017	-.0148	.0061	-.0153	-.0166	.0000	-.0654	-.0430	.0078	.0027
.3383	-.0045	-.0016	.0066	-.0007	-.3399	.2329	.2592		

TABLE L

COVARIANCE MATRIX FOR AGE GROUP 45-49

ROW NO. 1									
5.4081	25.7588	-.0039	-.0060	.0194	.0220	.0572	.0075	.0466	-.0271
-.0134	-.0384	-.CC24	-.C504	.0856	.0273	-.0596	-.0136	.0218	.3301
1.5944	-.0115	-.0164	.0149	.0096	4.2400	.7699	.5600		
ROW NO. 2									
25.7588	580.0198	2.4955	-.0712	.8889	.9602	4.0092	4.2328	2.8719	-.1484
.3480	-.0138	.3726	2.5454	1.8285	.5122	-.4685	1.5010	-.0600	15.8642
20.7172	.2283	-.0835	.2900	.0524	-140.2943	2.9598	2.7659		
ROW NO. 3									
-.0039	2.4955	.0142	-.0002	.0045	.0049	.0219	.0250	.0157	-.0003
.0023	.0012	.CC23	.C152	.0122	.0021	-.0003	.0094	-.0009	.0846
.0768	.0017	-.0000	.0014	.0001	-.9225	-.0036	.0004		
ROW NO. 4									
-.0060	-.0712	-.0002	.0159	.0032	.0027	-.0011	-.0002	.0602	.0052
.0013	.0027	.CC06	.CC89	-.0018	-.0006	.0119	.0007	.0009	-.0197
-.0112	-.0005	.0144	-.0030	-.0002	.0504	-.0133	-.0086		
ROW NO. 5									
.0194	.8889	.0045	.0032	.0183	.0146	.0059	.0081	.0239	.0003
.0025	.0014	.CC06	.CC83	.0227	.0021	.0174	.0029	-.0005	.0413
.0330	.0009	.0058	-.0012	.0003	-.9372	-.0016	.0032		
ROW NO. 6									
.0220	.9602	.0049	.0027	.0146	.0202	.0058	.0092	.0443	-.0008
.0004	.0004	-.CC01	.CC61	.0400	.0020	.0193	.0035	-.0000	.0440
.0422	.0009	.0048	-.0011	.0001	-.9000	-.0058	.0005		
ROW NO. 7									
.0572	4.0092	.0219	-.0011	.0059	.0058	.2335	.1001	-.0023	-.0025
.0051	.0060	.CC68	.CC73	-.0386	.0040	-.0014	.0330	-.0012	.0801
.1057	.0011	.0065	.0047	.0005	-2.2507	-.0049	-.0007		
ROW NO. 8									
.0075	4.2328	.0250	-.0002	.0081	.0092	.1001	.1492	.1214	-.0029
.0095	.0027	.CC89	.CC42	.0580	.0024	-.0085	.0207	-.0057	.1760
.0790	.0046	.0024	.0018	.0009	-1.9048	-.0256	-.0102		
ROW NO. 9									
.0466	2.8719	.0157	.0602	.0239	.0443	-.0023	.1214	6.2833	.1292
.0159	.0491	.C104	.1374	-.0210	-.0183	.0777	-.0056	-.0014	.0803
.3483	-.0018	.0703	-.0341	.0002	-1.7877	-.1938	-.1268		
ROW NO. 10									
-.0271	-.1484	-.0003	.0052	.0003	-.0008	-.0025	-.0029	.1292	.0423
.0021	.0047	.CC09	.CC77	-.0351	-.0006	.0058	-.0096	-.0007	-.0003
.0040	-.0001	.0064	-.0031	-.0004	.4231	-.0145	-.0156		
ROW NO. 11									

TABLE 2 (continued)

-.0134	.3480	.0023	.0013	.0025	.0004	.0051	.0095	.0159	.0021
.0201	.0026	.CC54	.C143	-.0217	.0023	-.0054	-.0008	.0006	.0297
.0345	.0006	.0019	-.0004	.0006	-.2167	-.0032	.0006		

ROW NO. 12

-.0384	-.0138	.0012	.0027	.0014	.0004	.0060	.0027	.0491	.0047
.0026	.0293	.CC50	.C17C	-.0305	-.0001	.0071	.0046	-.0009	-.0079
-.0231	-.0000	.0028	-.0005	-.0006	.1067	-.0200	-.0196		

ROW NO. 13

-.0024	.3726	.0023	.0006	.0006	-.0001	.0068	.0089	.0104	.0009
.0054	.0050	.C481	.CC72	.0169	.0010	-.0194	.0043	-.0006	.0238
.0217	.0003	.0000	.0007	-.0008	.0318	-.0042	-.0090		

ROW NO. 14

-.0504	2.5454	.0152	.0089	.0083	.0061	.0073	.0042	.1374	.0077
.0143	.0170	.CC72	.2886	.0197	.0080	-.0007	.0151	-.0077	.1102
.1312	.0019	.0107	-.0040	.0013	-.1520	-.0576	-.0359		

ROW NO. 15

.0856	1.8285	.0122	-.0018	.0227	.0400	-.0386	.0580	-.0210	-.0351
-.0217	-.0305	.C169	.C197	3.0474	-.0212	.5214	-.0431	-.0111	.1867
-.0349	.0064	-.0044	.0136	.0049	-1.0603	-.0545	.0218		

ROW NO. 16

.0273	.5122	.0021	-.0006	.0021	.0020	.0040	.0024	-.0183	-.0006
.0023	-.0001	.CC1C	.CC80	-.0212	.0512	.0342	.0031	.0010	.0236
.0246	.0003	.0002	-.0003	.0002	-.4042	.0017	.0041		

ROW NO. 17

-.0596	-.4685	-.0003	.0119	.0174	.0193	-.0014	-.0085	.0777	.0058
-.0054	.0071	-.C194	-.CC07	.5214	.0342	1.5262	.0830	.0086	-.0482
-.0291	-.0004	.0172	-.0154	.0048	-.1686	-.0108	.0022		

ROW NO. 18

-.0136	1.5010	.0094	.0007	.0029	.0035	.0330	.0207	-.0056	-.0096
-.0008	.0046	.CC43	.C151	-.0431	.0031	.0830	.3547	-.0111	.0861
.0319	.0023	.0015	-.0033	.0017	-.3481	-.0113	.0083		

ROW NO. 19

.0218	-.0600	-.0009	.0009	-.0005	-.0000	-.0012	-.0057	-.0014	-.0007
.0006	-.0009	-.CC06	-.CC77	-.0111	.C010	.0086	-.0111	.0278	-.0019
.0167	-.0003	.0009	.0009	.0002	-.0123	.0030	.0048		

ROW NO. 20

.3301	15.8642	.0846	-.0197	.0413	.0440	.0801	.1760	.0803	-.0003
.0297	-.0079	.C238	.1102	.1867	.0236	-.0482	.0861	-.0019	1.7163
1.1393	.0395	-.0228	.0277	.0018	-8.5066	-.0209	.0570		

ROW NO. 21

1.5944	20.7172	.0768	-.0112	.0330	.0422	.1057	.0790	.3483	.0040
.0345	-.0231	.0217	.1312	-.0349	.0246	-.0291	.0319	.0167	1.1393
3.0054	-.0041	-.0134	.0181	.0044	-4.2114	.4776	.3561		

ROW NO. 22

-.0115	.2283	.0017	-.0005	.0009	.0009	.0011	.0046	-.0018	-.0001

TABLE L (continued)

.0006	-.0000	.0003	.0019	.0064	.0003	-.0004	.0023	-.0003	.0395
-.0041	.0013	-.0005	.0006	.0000	-.2200	-.0072	-.0030		
ROW NO. 23									
-.0164	-.0835	-.0000	.0194	.0058	.0048	.0065	.0029	.0703	.0064
.0019	.0028	.0000	.0107	-.0044	.0002	.0172	.0015	.0009	-.0228
-.0134	-.0005	.0249	-.0039	-.0002	.1630	-.0155	-.0054		
ROW NO. 24									
.0149	.2900	.0014	-.0030	-.0012	-.0011	.0047	.0018	-.0341	-.0031
-.0004	-.0005	.0007	-.0040	.0136	-.0003	-.0154	-.0033	.0009	.0277
.0181	.0006	-.0039	.0167	-.0003	-.3428	.0034	.0052		
ROW NO. 25									
.0096	.0524	.0001	-.0002	.0003	.0001	.0005	.0009	.0002	-.0004
.0006	-.0006	-.0008	.0013	.0049	.0002	.0048	.0017	.0002	.0018
.0044	.0000	-.0002	-.0003	.0027	-.1044	.0046	.0051		
ROW NO. 26									
4.2400	-140.2943	-.9225	.0504	-.9372	-.9000	-2.2507	-1.9048	-1.7877	.4231
-.2167	.1067	.0318	-.1520	-1.0603	-.4042	-.1686	-.3481	-.0123	-8.5066
-4.2114	-.2200	.1630	-.3428	-.1044	981.8854	2.5652	.9339		
ROW NO. 27									
.7699	2.9598	-.0036	-.0133	-.0016	-.0058	-.0049	-.0256	-.1938	-.0145
-.0032	-.0200	-.0042	-.0576	-.0545	.0017	-.0108	-.0113	.0030	-.0209
.4776	-.0072	-.0155	.0034	.0046	2.5652	.4346	.2840		
ROW NO. 28									
.5600	2.7659	.0004	-.0086	.0032	.0005	-.0007	-.0102	-.1268	-.0156
.0006	-.0196	-.0090	-.0359	.0218	.0041	.0022	.0083	.0048	.0570
.3561	-.0030	-.0054	.0052	.0051	.9339	.2840	.3174		

TABLE LI

COVARIANCE MATRIX FOR AGE GROUP 50-64

ROW NO.	1									
	5.5585	25.3501	-.0126	-.0130	.0021	.0083	-.0105	.0407	.0356	.0071
	-.0180	-.0136	-.0079	-.0440	.1107	-.0053	.1501	.0016	.0026	.3482
	1.7395	-.0139	-.0194	.0182	.0173	6.8725	.5970	.4594		
ROW NO.	2									
	25.3501	572.0039	2.4157	.0139	.6582	.7330	3.6689	4.5878	.8632	-.2238
	.3160	-.0129	.4866	1.7711	3.2736	-.0454	1.3256	1.0437	-.0224	17.1220
	26.8487	.1848	-.2515	.1023	.1665	-150.0113	1.0304	2.0022		
ROW NO.	3									
	-.0126	2.4157	.0139	.0003	.0040	.0041	.0210	.0250	.0029	-.0014
	.0023	.0004	.0029	.0110	.0152	.0001	.0013	.0052	-.0001	.0860
	.0995	.0014	-.0009	.0004	.0004	-1.0795	-.0104	-.0007		
ROW NO.	4									
	-.0130	.0139	.0003	.0157	.0034	.0037	.0024	.0024	.0487	.0052
	.0008	.0017	.0012	.0047	-.0047	.0014	.0018	.0050	.0004	.0118
	.0091	.0002	.0139	-.0023	-.0002	-.0590	-.0067	-.0070		
ROW NO.	5									
	.0021	.6982	.0040	.0034	.0203	.0145	.0006	.0058	.0219	.0011
	.0014	-.0008	.0004	.0055	.0351	.0003	-.0073	-.0058	.0015	.0510
	.0327	.0012	.0052	.0001	.0007	-.4879	-.0095	-.0059		
ROW NO.	6									
	.0083	.7330	.0041	.0037	.0145	.0191	.0019	.0079	.0372	.0005
	.0007	-.0012	-.0004	.0049	.0368	-.0002	-.0018	-.0040	.0012	.0418
	.0364	.0008	.0045	.0011	.0008	-.6038	-.0033	-.0003		
ROW NO.	7									
	-.0105	3.6689	.0210	.0024	.0006	.0019	.2193	.0945	-.0084	-.0036
	.0097	.0046	.0210	.0378	-.0133	.0003	-.0321	-.0009	.0081	.0668
	.0791	.0012	.0027	.0004	.0020	-1.7583	-.0144	.0251		
ROW NO.	8									
	.0407	4.5878	.0250	.0024	.0058	.0079	.0945	.1417	.0416	-.0056
	.0070	-.0004	.0044	.0161	.0163	.0011	.0115	-.0016	.0004	.1617
	.1275	.0035	.0006	.0010	.0010	-2.4824	-.0377	-.0097		
ROW NO.	9									
	.0356	.8632	.0029	.0487	.0219	.0372	-.0084	.0416	6.0301	.1232
	.0187	.0468	.0377	.0790	-.2493	.0025	.0297	.0005	.0257	-.1279
	.3307	-.0093	.0578	.0005	-.0063	3.3995	-.0016	-.0177		
ROW NO.	10									
	.0071	-.2238	-.0014	.0052	.0011	.0005	-.0036	-.0056	.1232	.0438
	.0023	.0039	-.0002	-.0034	-.0800	.0011	-.0028	-.0034	.0030	-.0096
	.0056	-.0004	.0071	-.0011	-.0013	.3152	-.0018	-.0059		
ROW NO.	11									

TABLE B1 (Continued)

-.0180	.3160	.0023	.0008	.0014	.0007	.0097	.0070	.0187	.0023
.0219	.0036	.0046	.0130	-.0216	-.0011	-.0222	.0030	.0001	.0240
.0117	.0006	.0018	-.0020	.0003	-.2346	-.0075	-.0058		
ROW NO. 12									
-.0136	-.0129	.0004	.0017	-.0008	-.0012	.0046	-.0004	.0468	.0039
.0036	.0270	.0074	.0085	-.0717	-.0011	-.0281	.0006	-.0001	.0173
-.0040	.0006	.0020	-.0013	.0007	-.1158	-.0082	-.0132		
ROW NO. 13									
-.0079	.4866	.0029	.0012	.0004	-.0004	.0210	.0044	.0377	-.0002
.0046	.0074	.0468	.0155	.0014	.0010	-.0117	.0021	.0003	.0224
.0153	.0005	.0010	-.0011	.0005	-.3547	-.0074	-.0084		
ROW NO. 14									
-.0440	1.7711	.0110	.0047	.0055	.0049	.0374	.0161	.0790	-.0034
.0130	.0089	.0155	.2530	.0406	-.0060	.0102	.0124	-.0044	.0416
.0507	.0008	.0078	.0024	.0013	-1.2420	.0294	.0317		
ROW NO. 15									
.1107	3.2736	.0152	-.0047	.0351	.0368	-.0133	.0163	-.2493	-.0800
-.0216	-.0717	.0014	.0406	2.9234	.0037	.6356	-.0638	-.0203	.0202
.0423	-.0000	-.0109	.0091	-.0001	.5219	.0986	.0867		
ROW NO. 16									
-.0053	-.0454	.0001	.0014	.0003	-.0002	.0003	.0011	.0025	.0011
-.0011	-.0011	.0010	-.0066	.0037	.0769	.0238	-.0013	-.0025	-.0016
-.0097	.0000	.0000	-.0002	-.0012	-.6289	-.0087	-.0094		
ROW NO. 17									
.1501	1.3256	.0013	.0018	-.0073	-.0018	-.0321	.0115	.0297	-.0028
-.0222	-.0281	-.0117	.0102	.6356	.0238	1.4326	.0959	-.0057	.0145
.0265	.0004	.0042	-.0032	.0002	-.8327	.0259	.0323		
ROW NO. 18									
.0016	1.0437	.0052	.0050	-.0058	-.0040	-.0009	-.0016	.0005	-.0039
.0030	.0006	.0021	.0124	-.0638	-.0013	.0959	.3456	-.0175	.0335
.0447	.0006	.0028	-.0021	.0007	-2.2461	-.0100	.0011		
ROW NO. 19									
.0026	-.0224	-.0001	.0004	.0015	.0012	.0081	.0004	.0257	.0030
.0001	-.0001	.0003	-.0044	-.0203	-.0025	-.0057	-.0175	.0268	.0090
.0138	.0002	.0017	.0008	.0004	.1233	-.0095	-.0114		
ROW NO. 20									
.3482	17.1220	.0860	.0118	.0510	.0418	.0668	.1617	-.1279	-.0096
.0240	.0173	.0224	.0416	.0202	-.0016	.0145	.0335	.0090	1.1773
1.1291	.0409	-.0036	.0070	.0096	-9.9244	-.0749	.0108		
ROW NO. 21									
1.7395	26.8487	.0995	.0091	.0327	.0364	.0791	.1275	.3307	.0056
.0117	-.0040	.0153	.0507	.0423	-.0097	.0265	.0447	.0138	1.1291
3.3042	-.0087	-.0091	.0089	.0150	-5.4259	.4658	.3386		
ROW NO. 22									
-.0139	.1848	.0014	.0002	.0012	.0008	.0012	.0035	-.0093	-.0004

TABLE LI (continued)

	.0006	.0006	.0005	.0008	-.0000	.0000	.0004	.0006	.0002	.0009
	-.0087	.0014	.0000	.0001	.0001	-.2604	-.0090	-.0050		
ROW NO. 23										
	-.0194	-.2515	-.0009	.0139	.0052	.0049	.0027	.0008	.0578	.0071
	.0018	.0020	.0010	.0078	-.0109	.0000	.0042	.0028	.0017	-.0036
	-.0091	.0000	.0250	-.0035	-.0003	.2394	-.0017	-.0015		
ROW NO. 24										
	.0182	.1023	.0004	-.0023	.0001	.0011	.0004	.0010	.0005	-.0011
	-.0020	-.0013	-.0011	.0024	.0091	-.0002	-.0032	-.0021	.0008	.0070
	.0089	.0001	-.0035	.0167	.0003	-.4881	.0027	.0104		
ROW NO. 25										
	.0173	.1665	.0004	-.0002	.0007	.0008	.0020	.0010	-.0063	-.0013
	.0003	.0007	.0005	.0013	-.0001	-.0012	.0002	.0007	.0004	.0096
	.0150	.0001	-.0003	.0003	.0038	.0632	-.0004	-.0020		
ROW NO. 26										
	6.8725	-150.0113	-1.0795	-.0590	-.4879	-.6038	-1.7583	-2.4824	3.3995	.3152
	-.2346	-.1158	-.3547	-1.2420	.5219	-.6289	-.8327	-2.2461	.1233	-9.9244
	-5.4259	-.2604	.2394	-.4881	.0632	1030.6964	1.0549	-.8520		
ROW NO. 27										
	.5970	1.0304	-.0104	-.0067	-.0095	-.0033	-.0144	-.0377	-.0016	-.0018
	-.0075	-.0082	-.0074	.0294	.0986	-.0087	.0259	-.0100	-.0095	-.0749
	.4658	-.0090	-.0017	.0027	-.0004	1.0549	.4071	.2751		
ROW NO. 28										
	.4594	2.0022	-.0007	-.0070	-.0059	-.0003	.0251	-.0097	-.0177	-.0059
	-.0058	-.0132	-.0089	.0317	.0867	-.0094	.0323	.0011	-.0114	.0108
	.3386	-.0050	-.0015	.0104	-.0020	-.8520	.2751	.3398		

VITA

Bernard J. Schroer

Candidate for the Degree of

Doctor of Philosophy

Thesis: AN AUTOMATED TECHNIQUE FOR PATIENT HEALTH ANALYSIS

Major Field: Engineering

Biographical:

Personal Data: Born October 11, 1941, in Seymour, Indiana, the son of Alvin John and Selma Ann Schroer, of Seymour, Indiana.

Education: Attended high school in Seymour, Indiana, and graduated in 1959. Entered Purdue University in 1959, and received an Associate in Mechanical Technology in 1961. Entered Western Michigan University in 1962, and received the Bachelor of Science in Engineering degree in 1964. Entered the University of Alabama in 1964, and received the Master of Science in Engineering degree in 1967. Entered Oklahoma State University in 1970, and completed requirements for the Doctor of Philosophy degree in May, 1972.

Professional Experience: Employed by Sandia Corporation, Albuquerque, New Mexico, from June 1961, to September, 1962, as a Mechanical Designer. Employed by Brown Engineering Company, Huntsville, Alabama, from June, 1964, to September, 1967, as a Senior Engineer. Employed by The Boeing Company, Huntsville, Alabama, from September, 1967, to April, 1970, as a Research Engineer. Employed by Computer Sciences Corporation, Huntsville, Alabama, from April, 1970, to September, 1970, as a Project Engineer. Employed by Oklahoma State

University, Stillwater, Oklahoma, from September, 1970, to June, 1971, as a Graduate Assistant in the School of Industrial Engineering and Management. Presently employed with Computer Sciences Corporation, Huntsville, Alabama, as a Project Engineer.

Professional Membership: American Institute of Industrial Engineers, Association for Computing Machinery, and Registered Professional Engineer.