HUMAN PATHOLOGY

A Clinicopathologic Quarterly

THE ANALYSIS OF CLINICAL LABORATORY DATA

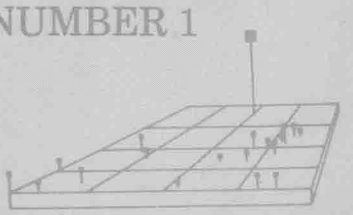
Ellis S. Benson and Howard M. Rawnsley, *Guest Editors*

March 1973

PUBLISHED BY W. B. SAUNDERS COMPANY

IN THIS ISSUE

VOLUME 4
NUMBER 1



STATEMENT OF PURPOSE

Human Pathology is designed to bring to the laboratory and clinical physician, and to the student, authoritative information of clinicopathologic significance to human disease. Its primary goal is the presentation of information drawn from morphologic and clinical laboratory studies having direct relevance to the understanding of diseases of man. Accordingly, it is intended that the journal embrace articles concerned with morphologic and clinicopathologic observations, reviews of a disease or group of diseases, analyses of problems in pathology, significant collections of case material, and new advances in concepts or techniques of value in the analysis and diagnosis of disease. Theoretical and experimental pathology and molecular biology pertinent to the diseases of man will be included.

The journal articles will be of interest to all physicians engaged in the clinical and laboratory practice of medicine as well as to teachers of pathology and clinical medicine. They will be critical and authoritative. Above all, they will embody a point of view.

SUBSCRIPTION INFORMATION

Subscription rate, United States, $18.50 per year. Inquiries and subscription orders may be addressed to W. B. Saunders Company, West Washington Square, Philadelphia, Pennsylvania 19105; in Canada to W. B. Saunders Company, Canada, Ltd., Toronto 7, Ontario; and in England to W. B. Saunders Company, Ltd., 12 Dyott Street, London WC1A 1DB.

HUMAN PATHOLOGY
A Clinicopathologic Quarterly

VOLUME 4
NUMBER 1
March 1973

Current Topics

CONSOLIDATION AND REGIONALIZATION OF PATHOLOGY SERVICES

IAN J. L. GOLDBERG, PH.D., F.R.C. PATH.*

In the last five years extremely important advances have been made in laboratory medicine. If these are developed in an orderly way, they should insure a central role for laboratory sciences in the provision of health care. There are also great dangers that they may be abused, with serious consequences.

These advances are most clearly evident in clinical chemistry, but the same principles apply to microbiology, immunology, and hematology as well. Sophisticated (and expensive) physical and radiochemical instruments find a very receptive market in hospital laboratories, which are rapidly learning how to use them effectively. These laboratories are leaving behind a preoccupation with labor saving devices and are beginning to look at means for attaining accurate diagnosis and treatment at very low cost and in the shortest possible time. Fast and precise, computer controlled analyzers using discrete, continuous flow, or centrifugal modes of operation will enable reference values for most assays to be determined, allowing for postural, geographic, ethnic, economic, age, and sex differences. Methods for the accurate estimation of different hormones, trace elements, and important drugs and enzymes are opening new areas of laboratory medicine.

It is this information explosion, both quantitative and qualitative, that is creating both the opportunity and the hazard. Serious mistakes have already been made, such as the indiscriminate use of screening, dependence on research workers for routine tests, and attempts to do too much with inadequate resources. Existing technology can generate, calculate, and distribute up to 6000 chemical analyses every hour. It can also support at least six subspecialities of clinical chemistry alone (bulk or demographic analysis, hormones, enzymes, proteins, pharmacological and toxicological biochemistry, lipids). However, equipment capable of producing supplementary and precise information is very expensive, and analysts capable of doing well the complex tests needed when automation is not appropriate are also scarce and expensive.

The use of fast, controlled analyzers and the recognition of subspecialization make it necessary to determine the optimal population module for a clinical chemistry service. Studies in Britain have shown that some analyses are so rarely needed that regional or even national centers of excellence should be set up for them.[1, 2] A very nearly comprehensive service could be established for a population of 500,000 to 750,000. Because this population would usually be served by two or three large hospitals, or a greater number of smaller ones, cooperation is necessary. This should take place in a progressive way; for example, a start could be made by setting up a regional service for less common hormone or radioimmunoassay estimations and by gradually extending the range of cooperation when it has been demonstrated that specimen transport, staff training, and other problems have been solved. It has also occasionally proved possible to start by centralizing bulk analysis, at least to the extent of providing an admission "profile" for all patients admitted to the hospital from the population area.[3]

All admitting hospitals involved must of course retain fully comprehensive facilities for all "in-office-hours" work and should also retain a clinical chemistry presence of significant

*Consultant Chemical Pathologist, Department of Chemical Pathology, St. Mary's Hospital, London, England.

magnitude by having at least one major subspeciality developed at it. Since clinical chemistry is such an important growth area in present day health care, no hospital with emergency beds can hope to function effectively without such a facility. (This comment probably also applies to the other divisions of human pathology and when rationalization is contemplated, should be remembered.) These opportunities created by rapid technological advances should not be sacrificed because of parochial attitudes. Proficiency requirements are rightly becoming more stringent, and work continues to increase in volume and range. There is no real prospect that there will be a sudden diversion of high caliber scientists and doctors into laboratory medicine to improve standards and maintain the output. Even if this prospect did exist, the cost of "going it alone" is too high for the individual hospital.

REFERENCES

1. Hospital Scientific and Technical Services: Report of the Committee, 1967–1968. London, H. M. Stationery Office, 1968.
2. Goldberg, I. J. L., and Mitchell, F. L.: Clinical biochemistry in the hospital laboratory service. Lancet, *1*:1240, 1970.
3. Letellier, G.: Personal communication.

CLINICAL PATHOLOGY IN THE NEW INDIVIDUALIZED MEDICAL SCHOOL CURRICULUM

ELLIS S. BENSON, M.D.,*
G. MARY BRADLEY, M.D.,†
AND PATRICK C. J. WARD, M.D.‡

No consensus appears to have arisen concerning the place of teaching clinical pathology (laboratory medicine) to medical students in the United States. Some educators favor teaching this subject, or a collection of subjects, in association with anatomic pathology in the first or second year, often as part of a course in human pathophysiology. In other medical schools the topics loosely grouped under the heading "clinical pathology" are taught by the Department of Medicine within its introductory courses in the first two years.

Similarly no firm consensus has emerged concerning the method of teaching clinical pathology, especially the question of how great a part laboratory exercises should play and what their purpose should be. Emphasis has shifted from the position that physicians must be able to perform skillfully a variety of simple clinical-laboratory examinations to the realization that few of them have either the time or inclination to perform them, and most physicians delegate these tasks to others. At the same time the need to understand the complexities and limitations of laboratory tests and to analyze laboratory data perceptively has gained increased attention in the teaching of clinical pathology.

This lack of consensus, which undoubtedly reflects dissatisfaction with traditional methods of teaching laboratory medicine, should be a goad toward innovative approaches, for which the new so-called individualized curricula in medicine provide opportunities.

At the University of Minnesota we are experimenting with three types of teaching programs in laboratory medicine. The first of these is a practical introductory "minicourse" (four weeks) taught in the second year "core" curriculum as part of a program called "The Student as a Physician." This program aims to teach the student basic clinical skills preparatory to his experience as a clinical clerk. These include history taking, physical examination, and laboratory examination of the patient. Laboratory skills of specimen acquisition (venipuncture, blood culture), routine screening procedures, complete blood count, urinalysis, and sputum and spinal fluid examination are taught. Lectures emphasize the use and misuse of the laboratory, quality control, standards of reference, and sources of error. Laboratory analyses are correlated with brief case histories. This aspect is appreciated by the students, who are concurrently learning the acquisition of data by history and physical examination. In an effort to encourage the learning of skills and case discussion, a student-staff ratio of 8 to 1 is maintained.

The second method of instruction involves taking part in an integrated basic-science and clinical system-and-organ-oriented curriculum, which extends throughout the second year. This effort has been most successful in the teaching program on the hematopoietic system, the kidney, and fluid and electrolyte balance.

Experiences that have been particularly enlightening to both students and faculty have arisen from our presentation of topics in laboratory medicine in a series of elective courses in the third and fourth years, during which the curriculum is now entirely elective. We have organized electives in special topics in the divisions of hematology, clinical chemistry, clinical microbiology, blood coagulation, and clinical laboratory immunology. These are

*Department of Laboratory Medicine, University of Minnesota, Minneapolis, Minnesota.
†Department of Laboratory Medicine, University of Minnesota, Minneapolis, Minnesota.
‡Department of Laboratory Medicine, University of Minnesota and Mount Sinai Hospital, Minneapolis, Minnesota.

small group sessions with a close interaction between faculty and students. The emphasis is placed on clinical problem-solving using laboratory data analysis.

What has been most illuminating, to the faculty at least, has been the success of a broad-based teaching program on the "Interpretation of Laboratory Data" at an affiliated community hospital. This is offered as a six week elective limited to 25 students in each of three quarters of the senior year. The course has been fully subscribed each quarter. Three clinical pathologists teach four hours a day, using case material from their own laboratories, and the course covers problems of interpretation in chemistry, hematology, and immunology. The students are given advance reading assignments to prepare them for each day's discussions, which are lively. It is obvious that at this stage of their experience the students are able to enter actively into discussions of laboratory tests, their uses, and their interpretation. Students are now aware of unfulfilled potentials of laboratory methods and are eager to increase their ability to use the laboratory in clinical problem solving.

A byproduct of this program has been a new interest among medical students in pathology as a career. They come to regard the pathologist not only as a scientist but also as a physician with important laboratory skills who is able to assist in solving complex clinical problems by the use of laboratory data.

What has emerged from these experiences, in our minds at least, is the realization that in order to have maximal relevance to the medical school experience of the new physicians, clinical pathology should be taught not only early in the curriculum but also later when the student is acquiring his clinical competence through first hand experiences in patient care. The early teaching affords a background for his clinical clerkship learning effort; the later teaching places his laboratory ability in the context of his entire clinical experience and provides a basis for his optimal utilization of laboratory sciences and data as an intern and resident.

Ellis S. Benson, M.D., and Howard M. Rawnsley, M.D.,
Guest Editors

INTRODUCTION

The series of papers that make up this special issue of *Human Pathology* are all directed to some aspect or other of a topic that is increasingly engaging the attention and time of pathologists and other laboratory physicians and scientists, namely, the analysis of laboratory data. Improvements in instrumentation, automation, laboratory proficiency, specimen handling, and reporting have all made it possible to provide large quantities of data in relatively short periods of time. Unfortunately, much relevant data goes unused or underused because of inadequate methods of analysis. It is the responsibility of laboratory physicians and scientists to assist in the analysis of clinical laboratory data by providing improved and more powerful methods of analysis and the necessary background information. The laboratory data being accumulated for each patient should be used to its fullest extent, and all its informational content should become a part of each patient's record to be used appropriately by the physicians who take the responsibility for his care.

The papers composing this symposium represent beginning efforts to address either the problems of accumulating needed background information such as analytical variation, physiological variation, or genetic variation ("normal range") or the problems of analyzing the data in the context of the entire clinical picture or selected parts of it. The results of clinical laboratory analysis are seen as being applicable to at least three types of problems: screening for unsuspected disease, following up and commenting on diagnostic leads or clues, and monitoring the course of the disease and its management. Different types of background information and methods of analysis are needed in each of these tasks. Dr. Eugene Johnson, in an introductory article, describes and further defines these needs and appeals to laboratory physicians and scientists to attempt to meet them.

No very complicated or sophisticated attempts to analyze laboratory data are included in the present series. An illustration of the effectiveness of analyses of laboratory data as "profiles" as a teaching mechanism in anecdotal form is provided by Dr. Ward. The attention of readers is called to earlier attempts to analyze the effectiveness of various laboratory tests of liver function using statistical means.[1] Recently Grams et al.[2] have called attention again to the possibilities inherent in multivariate analysis of laboratory data. It seems certain that further efforts will be made to exploit the informational content of clinical laboratory data by improved methods of analysis.

Ellis S. Benson, M.D.
Howard M. Rawnsley, M.D.

REFERENCES

1. Zieve, L., and Hill, E.: An evaluation of factors influencing the discriminative effectiveness of a group of liver function tests. I. The ultilization of multiple measurements in medicine. III. Relative effectiveness of hepatic tests in cirrhosis. Gastroenterology, *28*:766, 785, 1955.
2. Grams, R. R., Johnson, E. A., and Benson, E. S.: Laboratory data analysis system. Sections I through VI. Amer. J. Clin. Path., *58*:177, 1972.

SOME BASIC CONSIDERATIONS OF THE NEEDS FOR IMPROVED CLINICAL LABORATORY DATA ANALYSES

*Eugene A. Johnson, Ph.D.**

Abstract

A plea and sustaining arguments are presented for the organization of a data gathering system for obtaining the necessary background information to make the interpretation of clinical laboratory results stronger and more logical. Laboratory data are used in the solution of three types of clinical problems: screening for unsuspected disease, monitoring the course and management of a disease process, and investigation of diagnostic clues to arrive at a specific disease diagnosis. Background information to improve the use of laboratory data in these three tasks is badly needed. Clinical laboratory directors, pathologists, and clinical chemists are in a unique position to assemble this information and exploit its use. Who else will do it?

Departments of laboratory medicine and pathology have traditional clinical functions. They provide the trained personnel and the specialized equipment to make measurements on demand. Research has been a part of this tradition, but it has been oriented primarily toward the improvement of measurement techniques and the development of new measurement techniques that are a part of their service functions.

Ever increasing service demands on clinical laboratories, increasing variety of measurements, and the complexity of newer measurements all focus more attention on the role of the clinical laboratory. Concerted efforts have been made to improve logistics, striving for faster through-put, increased accuracy, and less cost. Critics will continue to register complaints, and research and development efforts mentioned will continue to be directed at the causes of these complaints. In this way the logistical efficiency of the clinical laboratory will continue to improve.

Another important aspect of the clinical laboratory's mission should be receiving new attention. How are the data being used? How are the test results being interpreted? Are they worth the cost, considering the uses to which the results are put? Is the maximum information inherent in the test being extracted from each test result?

Laboratory tests are requested for a variety of reasons: (1) They may be used to screen for unsuspected illness. A battery of tests is ordered and the reported results are the potential determinants of diagnostic suspicion and subsequent follow-up.

*Professor of Biometry and Laboratory Medicine, University of Minnesota, Minneapolis, Minnesota.

(2) Tests may be used to monitor the management of a patient with an established diagnosis, being used in sequence to detect changes. The changes may be interpreted as a natural part of the resolution of the disease process, or a worsening of the disease process, or as a signal of developing complications. (3) Tests may be requistioned to contribute specific diagnostic information. The diagnosis being still open, it is hoped that the test results will clarify the issue. Usually a "yes, no" signal is sought under these circumstances. These fundamentally different uses for laboratory data require different kinds of background information.

Screening for Unsuspected Illness

The utilization of the laboratory for screening purposes automatically creates a demand for certain kinds of basic background information. Screening is usually confined to the question of normal vs. non-normal. The distribution of test results for normals must be known before abnormal results can be defined. There is ample evidence for different distributions resulting from variations in measurement techniques used at different laboratories. Interlaboratory differences may occur even when the same techniques are used. Thus, each laboratory will have to establish its own norms for each test. General screening has fallen into some disrepute because of the low pay-off in detection of unsuspected disease states when compared to the total cost and the problem of frequent false positive results.[1] More emphasis is being placed upon screening for specific diseases with specific batteries of tests among special high risk groups. The norms for these specific groups must be established.[2] The results of such group specific screening should be continuously evaluated to determine whether the yield is sufficiently high to justify continuance. Concepts such as sensitivity, specificity, and prevalence are of great importance in screening.[3] Since much of the screening is based upon multiple test results, the use of multivariate analytical procedures as a means of enriching the analyses must be considered.[4] In multitest analysis, as in screening, certain statistical problems in clearly separating normals from abnormals have been identified and discussed.[5,6] The likelihood that a normal result will be wrongly identified as abnormal increases on repeated examination.[6]

Monitoring Patient Care

The utilization of the laboratory for monitoring patient care automatically creates a demand for a different kind of background information. Monitoring implies the surveillance of a sequence of laboratory reports in looking for changes. Change implies variability. What are the sources of variability? The laboratory must occasionally duplicate determinations on the same specimen to determine laboratory error distributions (analytical variance).[7,8] The laboratory should also obtain repeated specimens from the same normal subjects to determine the magnitude of normal temporal variation (physiological variance).[8] Only when these two ordinary, nonpathological sources of variation are understood can a proper assessment be made concerning changes potentially arising from changes in the patient's condition. For some determinations the laboratory error variance and the normal temporal variation within subject variance are both larger than the between-subject variance (genetic variance of "normal range").[8]

Diagnosis of Disease

The utilization of the clinical laboratory in disease diagnosis requires a third kind of background information. Clinical case reports in the medical literature on patients with a given disease diagnosis provide the results of laboratory determinations on the patients. At the very best, typical clinical research literature provides estimates of the multivariate distribution of test values to be expected from patients with a specified diagnosis. What it does not reveal is the likelihood of a specific diagnosis given a certain pattern of laboratory determinations.

A physician faced with a diagnostic quandary usually has available the information concerning the history, physical examination, presenting symptoms, and general clinical conditions. He then orders a battery of tests. His question is a condi-

tional one, conditional on all the clinical information and conditional on the presented multivariate test results. What are the possible diagnoses and their relative likelihoods? This is the kind of information that the clinical laboratory is in a unique position to provide.[4] In order to provide this type of information, however, the laboratory must develop and use a data gathering system that permits the gathering of this conditional frequency information. This implies that the circumstances surrounding the request for laboratory tests must be recorded as well as the test results and the final diagnosis. Such a system requires the cooperation of practicing diagnosticians and professional laboratory personnel in its design and maintenance. There is a good prospect that a proper choice of test batteries will yield multivariate patterns pathognomonic of specific disease states and changes within them. The search for such diagnostically potent batteries and the proof of their worth depend upon an elaborate, cooperatively maintained data gathering mechanism.

The Problem Posed

The statements regarding the need for specific background information made in earlier paragraphs are not new. For each reference cited the reader will probably be able to think of several other similar references. There have been many papers on normal values. Almost every laboratory now has some form of quality control program and as a consequence knows something about analytical error distributions. The issue is whether laboratory directors, pathologists, clinical biochemists, and other clinical laboratory professional personnel are fully recognizing both their opportunities and their responsibilities to provide laboratory data that are more meaningful and that make clear to the clinical staff all the implications of the test results. Are they attempting to make their work clinically relevant or are they hiding in a nonclinical realm? The problem is serious. It concerns the logical foundation for interpreting laboratory results.

Screening, in order to have a logical foundation, requires a knowledge of the group specific norms. At the very least, factors such as age, sex, race, and state of activity must be taken into account. Such data are most conveniently obtained by departments of laboratory medicine and pathology. Who else will do it? The monitoring of patients, in order to have a logical foundation, requires a knowledge of the magnitude of variation to be expected from laboratory error and repeat measurements on the same normal subject. Such data are most conveniently obtained by departments of laboratory medicine and pathology. Who else will do it? A pattern of test results for a given patient should suggest diagnostic possibilities and their weighting if the tests were ordered for that purpose. This implies that the probability of the presence of one disease state is different from that of another, conditional on the total clinical picture. Departments of laboratory medicine and pathology are in a unique position to establish the cooperative data gathering mechanism to obtain such information.

These ideas have been expressed before. Serious introspection is required of professional medical measurement specialists in clinical laboratories. Then, given a sufficient appreciation for the need to expand the logical aspects of clinical laboratory service, these specialists must orient new research in the direction described and thus also educate users.[9, 10]

The data gathering mechanisms discussed must be well designed and maintained. They must be guided by trained physicians, scientists, and technologists. It is not enough to try to capture the data as a trivial spin-off from an already existing patient accounting system. This means budgeting for the necessary additional expenses. It means setting up a cooperative schedule with the users of the laboratory so that each requisition potentially carries information concerning the reason for the request and other pertinent patient pretest status information. It means that the reports going back to the user will carry more information; the content may vary depending upon the reason for the original request. It means that the research file established by the accumulation of these many requisitions must be routinely processed for evaluation.

The records suggested here would be consistent with the problem oriented

charting suggested by Weed.[11] The difference would be that this is a special purpose file, maintained by departments of laboratory medicine or pathology. It would be constantly updated and constantly evaluated so that the clinical laboratory reports would always reflect the latest evaluation of the accumulated experience.

Many obvious difficulties can be anticipated in initiating the necessary educational efforts and system design changes. The potential benefits to medicine justify these efforts.

REFERENCES

1. Wilson, J. M. G.: Current trends and problems in health screening. Eighth International Congress on Clinical Chemistry, Copenhagen, 1972. Scand. J. Clin. Lab. Invest., *29*(Suppl. 126):22.1, 1972.
2. Keating, F. R.: The relationship of age and sex to distribution values in healthy adults of serum calcium, inorganic phosphorus, magnesium, alkaline phosphatase, total proteins, albumin and blood urea. J. Lab. Clin. Med., *73*:825–834, 1969.
3. Wilson, J. M. G., and Jungner, G.: Principles and Practice of Screening for Disease. Geneva, World Health Organization, 1968.
4. Grams, R. R., Johnson, E. A., and Benson, E. S.: Laboratory data analysis system. Section III. Multivariate normality. Section IV. Multivariate diagnosis. Amer. J. Clin. Path., *28*:188–207, 1972.
5. Schoen, I., and Brooks, S. H.: Judgment based upon 95% confidence limits: a statistical dilemma involving multitest screening and proficiency testing of multiple specimens. Amer. J. Clin. Path., *53*:190–193, 1970.
6. Schoenberg, B. S.: The "abnormal" laboratory result: problems in interpreting laboratory data. Postgrad. Med., *47*:151–155, 1970.
7. Caraway, W. T.: Accuracy in clinical chemistry. Clin. Chem., *17*:63–71, 1971.
8. Cotlove, E., Harris, E. K., and Williams, G. E.: Biological and analytic components of variation in long term studies of serum constituents in normal subjects. Clin. Chem., *16*:1016–1032, 1970.
9. Astrup, P.: Role of clinical chemist in hospitals. Ninth International Congress on Clinical Chemistry, Copenhagen, 1972. Scand. J. Clin. Lab. Invest., *29*(Suppl. 126):21.1, 1972.
10. Seligson, D.: The future of laboratory medicine. Eighth International Congress on Clinical Chemistry, Copenhagen, 1972. Scand. J. Clin. Lab. Invest., *29*(Suppl. 126):22.4, 1972.
11. Weed, L. L.: Medical Records, Medical Education and Patient Care. Cleveland, The Press of Case Western University, 1967.

Biometry Division
University of Minnesota
Minneapolis, Minnesota 55455

THE POPULATION OF HEALTHY PERSONS AS A SOURCE OF REFERENCE INFORMATION

*Lila Elveback, Ph.D.**

Abstract

The role of information about the population of healthy persons in the interpretation and evaluation of laboratory test results is discussed, and problems concerned with obtaining this information are considered. It is suggested that the inclusion of the age-sex specific percentile in healthy persons in each laboratory report would greatly improve its value to the physician.

The question of what set of reference information is of optimal usefulness to the clinician in evaluating a laboratory result in a given patient has given rise to a great deal of discussion. Multivariate methods based on large computer files of clinical and laboratory information obtained by following patients over a period of time have been proposed.[1] It seems certain that such information systems, as they become operative, will be of great value. The discussion of optimal methods of use of such systems and the role of so-called computer diagnosis in general will continue. It will be many years before the need for other and far simpler information sets will disappear.

Currently the reference standard most widely used is the distribution of values in healthy persons. The information concerning this distribution that is given to the physician usually consists of estimates of two percentiles of the distribution, most commonly, 2.5 and 97.5 points. This pair of numbers defines what has been called "the normal range." The inadequacy of this approach is the principal topic of this discussion.

The discussion that follows is limited to the choice of reference standards and the form of the report for the univariate problem. The importance of components of variation associated with laboratory measurement error, long and short term variation within the individual, seasonal and geographic differences, and the multivariate problem is discussed elsewhere. Only a single laboratory is considered, and it is assumed that all tests are performed with local residents and that they have been done at the same time of day under appropriately controlled conditions.

In the evaluation of current practice or of a proposed change, one useful step is to look at what is, and what can be achieved, from the viewpoint of the ideal. In evaluating a laboratory test result for a patient and reaching a decision as to whether immediate intervention or future careful surveillance is indicated, one property of

*Professor of Biostatistics, Mayo Graduate School of Medicine (University of Minnesota). Consultant, Department of Medical Statistics, Mayo Clinic and Mayo Foundation, Rochester, Minnesota.

the ideal set of reference data would be that it contain information regarding the prognostic significance of the result.

Ideally the information concerning prognostic significance would be complete enough so that information concerning the distribution in the currently healthy population, although it might still be of major public health importance, would be irrelevant to the care of the individual patient. The information might take the form of a table or graphs showing the risk (relative frequency of morbidity in the next k years of experience) for subgroups of persons by age, sex, level for the test in question, and other appropriate variables.[2-4] Although some prognostic information is available for a variety of tests, it is not sufficiently complete to permit the degree of quantification that would be necessary if it were to stand alone. It has been shown, for example, that elevations in cholesterol are associated with an increased risk of ischemic heart disease,[5] that some increase in the risk holds over the entire range of values, and that the extent of the increase in the risk with the cholesterol level may be age specific. The general conclusion, "the lower the better," is not sufficient for the purpose at hand, and further information is required. The distribution in the population of "healthy persons" is the first information requested. Cross sectional studies of this population, which are composed of the healthy-at-the-moment, will not add to our stock of prognostic information. The results are intended as a supplement to, or a substitute for, such information. Information concerning the distributions under a wide variety of alternatives to health is also desired, but is infrequently available except in broad qualitative terms.

Every self-respecting laboratory has some information that is related somehow to the distribution in the healthy. The remainder of this discussion concerns what this information should be and how it should be used.

It seems reasonable to begin by defining what constitutes an adequate set of information and an optimal way of using it, and to go on to explore the difficulties involved in obtaining the information.

1. The availability of an adequate number of healthy males and females with representation over the age range under study.

2. Test results, by current laboratory methods on these persons.

3. The availability of adequate statistical review during the collection of the data so that the sample size required and appropriate methods of estimation for the test in question can be determined.

4. A carefully thought out plan, resulting from close communication between the laboratory and the physician, for presenting the results of the test to the physician in a way that will minimize the demands on his time and memory and at the same time maximize the clinical application and usefulness.

Most institutions that could fulfill all these requirements would count themselves lucky indeed. However, to those about to set out to achieve them, the usefulness of the list is dependent on definitions, at least in the practical operational sense, of some of the vague terms and expressions.

Who Are the Healthy Persons?

Benson[6] has pointed out the difficulties involved with a demand for perfection. An operational definition that can yield useful results is all that is needed. No such definition is offered here, since this is obviously not a job for a statistician to tackle. However, I am convinced that a useful operational definition can be found. In the search for an optimally useful definition, there are several questions that should be asked.

1. Should the definition be one that will serve for all ages 20 and over, or should different definitions be used for several age groups?

2. Should a single broad definition be used that is appropriate for all laboratory tests, or should a basic definition be used that is designed to be supplemented as appropriate for each specific test?

3. Which persons should we accept? Which persons to disqualify (e.g., cancer, diabetes) is only one side of the question. Is the mere absence of a recorded diagnosis enough? Should an apparently healthy blood donor be accepted on the grounds that his medical record shows that he would have qualified three years

ago? Should we ask for some positive evidence of absence of disease? Keating and associates[7] required an essentially negative physical examination and normal findings on urinalysis, hemoglobin determination, leukocyte count, and stereoscopic roentgenograms of the heart and lungs for admission to their healthy group.

In what follows it will be assumed that some protocol has been adopted and that the word healthy denotes those who satisfy the requirements of this protocol.

The Availability of Healthy Persons

How are they to be identified? How are the test values to be obtained?

Ideally each institution would have a panel of healthy persons in the community, subject to periodic health examinations and prepared to appear on demand for the required tests. I do not know of such a panel, although it would seem that with time, money, and effort it should be possible to achieve.

In certain clinic situations the identification of an adequate number of healthy persons is possible on a continuing basis. It is also true that such identification is most frequently established after the sought-after individual has returned to the population at large. One method that has been used is to run the battery of tests in question on all persons in some subgroup, such as routine annual examinations, and subsequently select only those who qualify as healthy. The efficiency of the method depends on the subgroup chosen. I recall one study in which during the first months only about one quarter of the subgroup qualified and these were predominantly women under the age of 30.

The Mayo Clinic is in a particularly fortunate position, since it serves a large proportion of the local population and has an excellent record and retrieval system. Dr. Fred Nobrega and I have used this record system for a small study designed to determine the proportion of recent patients having elective surgery and blood donors who would qualify as healthy under a written protocol. Past experience indicated that of the routine annual examination group, only about 25 to 30 per cent were healthy. For the group identified through elective surgery the results were as follows (Table 1):

TABLE 1. "HEALTHY" PATIENTS IDENTIFIED THROUGH ELECTIVE SURGERY

Type of Surgery	*Number of Patients*	*Percentage Healthy*
Hemorrhoidectomy	207	45
Herniorrhaphy	100	36
Cataracts	100	32
Vein stripping	423	52

The majority of those disqualified had serious progressive disease. For example, 10 per cent had cancer, diabetes, or heart disease and another 10 per cent were hypertensive. In addition to the elective surgery groups we have also reviewed the histories of recent Mayo Clinic blood donors. We found that about three-fourths had been Clinic patients and that of these histories about 80 per cent were essentially negative. For a study of healthy persons including blood donors, adequate positive evidence of health might be taken to include an examination in the last 12 months. Of those with negative histories, we found that 54 per cent fulfilled this requirement. In the 232 histories examined we could qualify 43 per cent of the patients as healthy by our protocol. The ages and sex distributions of the healthy persons identified among blood donors and through elective surgery are given in Table 2.

For the small hospital the problem of identifying healthy persons is very difficult and the approach through elective surgery is frequently used. It is apparent that unless the hospital records contain sufficient information so that the "unhealthy" can be weeded out, the results may be highly misleading. The errors involved in mis-

TABLE 2. PERCENTAGE DISTRIBUTIONS BY AGE AND SEX

Age	*377 Patients Undergoing Elective Surgery*			*100 Healthy Blood Donors*		
	Male	*Female*	*Total*	*Male*	*Female*	*Total*
<40	6	16	22	42	22	64
40–59	19	30	49	28	7	35
60+	14	15	29	1		1
	39	61	100	71	29	100

classifying ill persons as healthy will in general be in the direction of giving too broad a normal range with resulting failure to detect some proportion of the patients in whom action, at least in terms of further testing, is appropriate. Some hospitals, realizing that their records do not allow adequate weeding out of the unhealthy, find false comfort in applying the Hoffman method to the results.[8] The inadequacies of this method are discussed in a later section.

Unfortunately no solution of this problem of the small hospital has been offered. The question of the extent to which the small laboratory can depend on a nearby large laboratory for assistance has been raised.[9] In such a system it is visualized that the laboratory of the largest hospital or medical center in a region would not only furnish the normal value studies and normal ranges derived therefrom, but would also of necessity serve as a standard of reference for all the smaller laboratories using these values. It seems clear that a study of the equipment and evaluation of performance in the smaller laboratories would be a necessary preliminary. Inherent in such an evaluation is a review of the available methodologies for all tests considered necessary for on-site performance in smaller facilities. A standard set of methodologies would then be selected for introduction into all participating peripheral laboratories in the region on the basis of their precision and how well the results compared with the results of the same tests generated in the central facility. This would be followed by the establishment of a common ongoing quality control program, which would include the use of the same standard and control materials. A continuing education program for laboratory personnel would be another assurance against poor performance. By taking these steps it may be possible to calibrate the results obtained by the small laboratory against those of the central facility and to maintain calibration by continued surveillance and education. The end result desired is to establish some set of tests that the small laboratory can employ for their patients—information concerning the distribution of values in healthy persons that has been accumulated by the central facility.

Is There a Sampling Problem?

If we agree that the healthy persons in the community are to serve as the reference population, do we need a representative sample? What is a representative sample?

Suffice it here to say that a sample will be called representative with respect to age if the age distribution of the sample agrees with the age distribution of the population, and representative with respect to age, sex, and race if the trivariate age-sex-race distribution of the sample agrees with that of the population.

In general there will be two sets of variables to consider. Set 1 includes age, sex, race, and any other variables chosen to define subgroups for which we want to establish the distribution in question. Set 2 includes those uncontrolled variables that influence the value in question (such as weight, diet, or occupation) but are excluded from consideration.

In these terms the question of representativeness can be pinned down. A sample "representative" with respect to the set 1 variables (age, race, and sex) may be far less efficient than one that gives equal numbers in the subsets (by race, for example) for which we need equally good estimates. For this set of variables we do not need, or even want, necessarily, representativeness. The use of regression on age (which is possible in the majority of cases) eliminates the necessity for subgroups by age and will in general give better age specific estimates. It is the uncontrolled set 2 variables that influence the level of the variable in question that needs to be of concern in this respect. It is a sad fact that it is seldom possible to study the representativeness, even though it is possible to collect the appropriate data on the sample members, since the distributions in the parent population are almost invariably unknown.

How many characteristics of the patient should be used in specifying the distribution against which his value is to be evaluated? I believe that it is widely accepted that any differences that exist in terms of age, sex, or race should be used; this would appear to be the minimal desirable set of variables. One of the great weaknesses of current practice is the fail-

ure to use this minimal set, even when the information is available.

The Hoffman Method

The Hoffman method for establishing information concerning the "healthy" sidesteps the entire problem of identifying healthy persons.[8] It is supposed to manufacture information about the distribution in healthy persons using two ingredients: the laboratory log of all tests done and a small amount of mathematics.

In order to accept the Hoffman method, it is necessary to accept the basic idea that information about "healthy persons" can be obtained without defining what is meant by this phrase or identifying a single person who deserves such a description. On closer examination of the method, it is found that it is necessary to accept three assumptions, each of which has been shown repeatedly to be false: (1) The distribution in healthy persons is gaussian (or "normal") in form. (2) The distribution in unhealthy persons is "gaussian" in form. (3) Mathematics alone will serve to separate out the healthy from whatever mixture of good and ill health is represented by the test in question.

The fact that distributions of physiologic variables cannot be assumed to be gaussian ("normal") in form has been recognized for at least 75 years. Both logic and experience show that the distributions in the highly heterogeneous group of the unhealthy are apt to be even more skewed than those in the healthy. Fourteen years have passed since Herrera[10] in the *Journal of Laboratory and Clinical Medicine* pointed out so clearly the implications of this fact for laboratory medicine. The literature over the last ten years contains countless illustrations of the nongaussian nature of these distributions in healthy persons. In the last five years a number of investigators have demonstrated the simple fact that the Hoffman method does not work.[11-15]

Further, it has been shown that the resulting errors can be of serious clinical significance, although this is not so in all cases. For example, for alkaline phosphatase the Hoffman method gave rise to misclassification as false negatives of 18 per cent of the laboratory values; for magnesium the error consisted of misclassification as false positives of only 1 per cent of the laboratory values; for serum iron the method resulted in a negative lower limit.[15] However, since in the application of the Hoffman method no hint is available as to whether the error is large or small, no reliance can be placed on the method.

What Is an Adequate Statistical Review?

An adequate statistical review is a review by an adequately trained (usually professional) statistician who has the time, and is willing to expend the effort: (1) to look at the data in terms of distribution form problems and possible transformations (including, but not stopping at, the logarithmic), (2) to consider the possibility of using regression on age, (3) to examine the form of the residuals about the regression function, (4) to decide about an appropriate method—parametric, nonparametric, or a combination of these—of estimating the percentiles, and (5) to consider whether the sample size is adequate in view of the distribution problem, the estimation problems, and the accuracy requirements associated with the test in question.

How Should the Values Be Reported to the Physician?

Progress in learning to measure, and to measure accurately, is one of the major components of progress in medicine. Each year millions of dollars are spent on research on measurement problems in the health sciences. The ultimate aim of such research is improvement in the health and medical care of the population, which must be implemented, at least in part, through improvements in the care of the individual patient. The accuracy with which determinations are made in the laboratory has improved greatly, with resulting benefit in research studies of many kinds. How has this improvement benefited the physician in the care of the individual patient? It is exactly at this point that the laboratory ignores the advantage of the measurement and allows the physician only a method of classification according to what Pickering[4] has called "the

fallacy of the dividing line." The "normal range," even as a classification system, is almost impossible to defend. It classifies 95 per cent of the healthy population as "usual" and 5 per cent as "unusual" for all patients and for all tests, quite without regard to the specific penalty for misclassification in any given application.

The "normal range" arises from putting together two very small facts concerning the healthy population with a large amount of wishful thinking. The result is a generalization, which brings to mind that "All generalizations are false, including this one." The finished product with its misnomer, "normal range," carries an implicit invitation to conclude that if a patient's value falls within the range he is healthy, and that if it lies outside the range he is not healthy and some action is needed. It is also implied that this same range will serve for the banker and the builder, the poet and the football star, the old and the young, the thin female and the obese male, the apparently well and those with serious concurrent disease.

Fortunately most clinicians avoid coming to such a conclusion and supplement the meager information the laboratory offers them in the "normal range" with experienced judgment based on knowledge of the individual patient and of the clinical implications of the test in question. Such experienced judgment will always play a role in clinical medicine. The report from the laboratory should be designed to be of maximal assistance in the formation of that judgment.

If the healthy population is to serve as the reference standard, some measurement of "how unusual in this population" is obviously preferable to any classification system. What are the criteria to be used in the selection of such a measurement? First, it is desirable that the scale be universal in the sense that any specified numerical value on the scale carries the same information on "how unusual" for all tests. Second, it should be such as to preserve this feature and at the same time allow an individual patient to be referred to whatever subgroup of the healthy population is appropriate for the test in question. As a minimum this subgroup should be specified in terms of age, race, and sex.

The suggestion that a common scale be adopted for laboratory results has appeared in the literature many times. The purpose of the common scale is to relieve the physician of the necessity of remembering or looking up a large number of reference values, and this is an important goal. The scale usually suggested has been the relative deviate (which tells how many standard deviations above or below the mean the value in question lies). This scale is universal in the sense that it is free of the units of measurement. It is not universal in the much more important sense that any specified value has the same relationship to relative frequency from one test to another. The relationship between the relative deviate and the relative frequency depends on the form of the distribution. The fact that this relationship is widely tabled for one particular form, the gaussian, is of no help in the majority of cases and has in fact resulted in a great deal of confusion and misinformation in the field of laboratory medicine. For example, in distributions based on several hundred healthy adults studied at the Mayo Clinic, the value of −2 on this scale corresponded for triglyceride to a value lower than any observed; for alkaline phosphatase 10 per cent had values lower than that corresponding to −2, whereas for serum iron the value corresponding to −2 was negative.[7, 15] As a measure of "how unusual in the healthy," the relative deviate is, to say the least, highly unsatisfactory. If the criterion for the selection of a common scale is that the values correspond in some constant way to "how unusual," it is difficult to see how any scale could have an advantage over that of the percentile.[16]

The percentile corresponding to any laboratory value is the percentage of the population in question that has this or lower values. For example, in Keating's study[7] the percentile corresponding to a blood urea of 37 mg. per 100 ml. for white males of age 65 is 51. It follows that 51 per cent of the healthy members of this subgroup have blood urea values this low. The percentile for white females of age 25 is 96, and it follows that only 4 per cent of the healthy persons in this group have blood urea levels higher than 37 mg. per 100 ml.

The existence of age and sex differences in levels is well known for many laboratory tests, and quantitative reports