Genetics and the epidemiology of chronic diseases : a symposium, June 17-19, 1963, Ann Arbor, Michigan ... / edited by James V. Neel, Margery Shaw, William J. Schull.

Contributors

Shaw, Margery W.
Schull, William J.
Neel, James V. 1915-2000.

Publication/Creation

Washington : U.S. Public Health Service, Division of Chronic Diseases, 1965.

Persistent URL

https://wellcomecollection.org/works/e3xr2j55

License and attribution

You have permission to make copies of this work under a Creative Commons, Attribution, Non-commercial license.

Non-commercial use includes private study, academic research, teaching, and other activities that are not primarily intended for, or directed towards, commercial advantage or private monetary compensation. See the Legal Code for further information.

Image source should be attributed as specified in the full catalogue record. If no source is given the image should be attributed to Wellcome Collection.

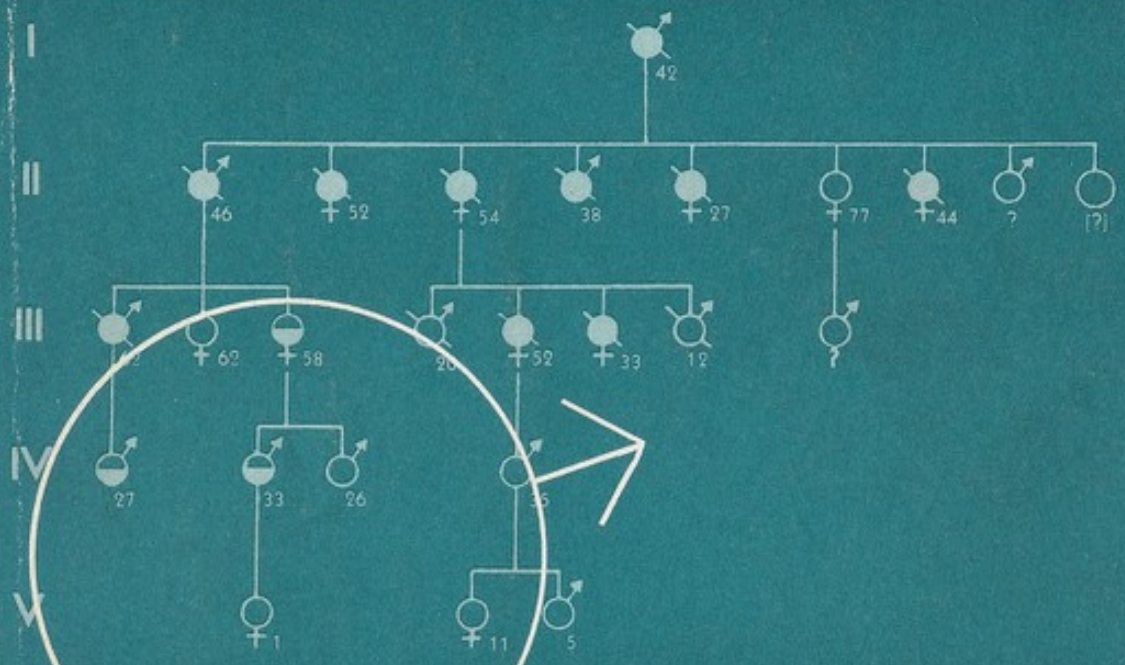
wellcome collection

Wellcome Collection
183 Euston Road
London NW1 2BE UK
T +44 (0)20 7611 8722
E library@wellcomecollection.org
https://wellcomecollection.org

THE ROYAL SOCIETY
for the Promotion
OF HEALTH
LIBRARY

28977

Genetics and the Epidemiology of Chronic Diseases



U.S. DEPARTMENT OF HEALTH,
EDUCATION, AND WELFARE
Public Health Service

SM/28

Public Health Service Publication No. 1163
February 1965

THE ROYAL SOCIETY

FOR THE PROMOTION

OF HEALTH

13, GROSVENOR PLACE, LONDON, SW1X 7EN.
~~90 Buckingham Palace Road, London, S.W.1~~

Class No. SM/28 Acc. No. 28977

This book is returnable on or before the last date Marked below

7 JAN 1966 18 12 79	

22900320099

X 155242

Genetics and the Epidemiology of Chronic Diseases

Edited by JAMES V. NEEL
MARGERY W. SHAW
WILLIAM J. SCHULL

Department of Human Genetics
University of Michigan Medical School
Ann Arbor, Michigan

A Symposium

June 17–19, 1963, Ann Arbor, Michigan,
Sponsored by U.S. Public Health Service,
University of Michigan Medical School,
University of Michigan School of Public Health

THE ROYAL SOCIETY
for the Promotion
OF HEALTH
LIBRARY

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Division of Chronic Diseases
Washington, D.C. 20201

1965

14480080

WELLCOME INSTITUTE LIBRARY
Coll. welMOmec
Call No. WT

Public Health Service Publication No. 1163

February 1965

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C., 20402 - Price $1.50 cents

Preface

The papers which follow were presented at a Symposium on Contributions of Genetics to Epidemiologic Studies of Chronic Diseases, held at Ann Arbor, Michigan, June 17–19, 1963, and co-sponsored by the Division of Chronic Diseases, U.S. Public Health Service, and the Schools of Medicine and Public Health of the University of Michigan.

The purpose of the symposium was to consider the role and limitations of the genetic approach in studies concerned with the epidemiology of a variety of chronic diseases. The papers presented have been organized under five headings, as follows: 1) Basic genetic and epidemiologic principles; 2) Critical reviews of selected, illustrative problems; 3) Problems in experimental design; 4) Papers describing current genetico-epidemiologic studies; and 5) Genetic counseling.

The first three groups are self-explanatory, but the fourth and fifth require a word of comment. The papers comprising the fourth group were delivered in a series of workshops in which the emphasis was as much on methodology and the limitations inherent in certain approaches as on actual results. An effort was made to ensure that these contributions mirrored the types of problems with which the epidemiologist and his confreres of the 1960's are occupied. The paper on genetic counseling was included in response to an expressed need for consideration of a topic which is increasingly coming to the attention of state health organizations.

The arrangers of the program and the editors are most grateful for the assistance in the planning rendered by the Symposium Planning Committee, the valuable services rendered by the Department of Postgraduate Medicine of the University of Michigan Medical School in the conduct of the program, and gratefully acknowledge financial assistance in the publication of the symposium proceedings by the Division of Chronic Diseases, the National Cancer Institute, and the National Institute of Arthritis and Metabolic Diseases of the Public Health Service.

SYMPOSIUM PLANNING COMMITTEE

Dr. Margery W. Shaw,
University of Michigan
Dr. Merton Honeyman,
Connecticut State Health Department
Dr. Lee Schacht,
Minnesota State Health Department
Dr. Alice Chenoweth,
Children's Bureau, Welfare Administration, U.S. Department of Health, Education, and Welfare

PUBLIC HEALTH SERVICE
U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Dr. Simon Abrahams,
Division of Radiological Health
Dr. Wilfred David,
Division of Chronic Diseases
Dr. Samuel Fox,
Heart Disease Control Branch, Division of Chronic Diseases
Dr. Glen McDonald,
Diabetes and Arthritis Branch, Division of Chronic Diseases
Dr. George K. Tokuhata,
Division of Chronic Diseases (Chairman)
Dr. Katherine S. Wilson,
National Institutes of Health
Dr. Carl Witkop, Jr.,
National Institutes of Health

Table of Contents

PART I

	Page
PREFACE	III
BASIC GENETIC AND EPIDEMIOLOGICAL PRINCIPLES	
Genetics and Epidemiology, by Thomas Francis, Jr	1
Genetic Studies of Family Units, by Bertram L. Hanna	7
Problems of Ascertainment in the Analysis of Family Data, by James F. Crow	23
Estimation of Genetic Parameters in Population Studies, by William J. Schull	45
Hereditary Factors Elucidated by Twin Studies, by Bent Harvald and M. Hauge	61
ABO Blood Groups, Secretor Status, and Susceptibility to Chronic Diseases: An Example of a Genetic Basis for Family Predispositions, by J. A. Fraser Roberts	77
Epidemiological Concepts Applied to Studies of Chronic Diseases, by Abraham M. Lilienfeld	87

PART II

THE EVALUATION OF GENETIC FACTORS IN SELECTED ILLUSTRATIVE DISEASES	
Diabetes Mellitus, by James V. Neel, Stefan S. Fajans, Jerome W. Conn, and Ruth T. Davidson	105
Coronary Artery Disease, by Victor A. McKusick	133
The Geographic and Genetic Characteristics of Amyotrophic Lateral Sclerosis and Other Selected Chronic Neurological Diseases, by Leonard T. Kurland	145

PART III

PROBLEMS IN THE EXPERIMENTAL DESIGN OF EPIDEMIOLOGICAL STUDIES	
Selection Techniques for Rare Traits, by Leslie Kish	165
The Reliability of Survey Data, by Charles F. Cannell	177
A Trial of Methods for Collecting Household Morbidity Data, by Kenneth R. Wilcox, Jr	185

PART IV

Page

CURRENT EPIDEMIOLOGICAL PROBLEMS

Intrafamilial Transmission of Rheumatoid Arthritis, by Sidney Cobb 209

Family and Genetic Studies of Rheumatoid Arthritis and Rheumatoid Factor in Blackfeet Indians, by Thomas A. Burch, William M. O'Brien, and Joseph J. Bunim 227

LONGEVITY AND CORONARY ARTERY DISEASE

Family Patterns of Longevity and Mortality, by Bernice H. Cohen 237

Coronary Heart Disease in Relation to Blood Pressure and Cholesterol Levels in Population Studies, by Fredrick H. Epstein and Marcus O. Kjelsberg 265

Inherited Antigenic Differences in Serum Beta Lipoproteins in Relation to Coronary Artery Disease and Diabetes, by Baruch S. Blumberg, Barbara Mishell and Sam Visnich 279

REGIONAL VARIATIONS IN MORBIDITY AND MORTALITY

Regional Variations in the Incidence of Spina Bifida, by David Hewitt 295

Vital Record Incidence of Congenital Malformations in New York State, by A. M. Gittelsohn and S. Milham, Jr. 305

Methodological Problems in the Use of Standard Vital Statistical Data in the Study of Neonatal Mortality and Birth Weight, by Douglas Grahn 321

MALIGNANT DISEASES

Familial Factors in Lung Cancer and Smoking, by George K. Tokuhata 339

Do Malignancies Result from Diagnostic and Therapeutic Radiation? by T. Sterling, E. Saenger, and R. A. Seltzer 355

Congenital Malformations and Cancer in the Sibships of Leukemic Children, by Robert W. Miller 373

PART V

GENETIC COUNSELING

Some Aspects of Genetic Counseling, by Franz J. Kallmann 385

PART I

Basic Genetic and Epidemiological Principles

Genetics and Epidemiology
Opening Comments

By Thomas Francis, Jr., M.D., *Henry Sewall University Professor of Epidemiology, Chairman of the Department of Epidemiology, School of Public Health, and Professor of Epidemiology, Department of Pediatrics and Communicable Diseases, Medical School, University of Michigan, Ann Arbor, Mich.*

I would like to join Dr. Neel in welcoming you to Michigan. The excellent program he and his associates have arranged with you and for you gives promise of stimulation and effect. We are fortunate in our environment here that the close relationship between the Departments of Epidemiology and Human Genetics provides a constant opportunity for interchange in study and ideas. We follow the current trend toward desegregation in these fields and hope that the present program will be of general aid in this direction.

The type of opening comments with which I am most familiar are phrases such as pass, or I can open, or I'll open for a couple. This relates to a very complicated intellectual exercise which I can explain in a form of genetic vulgate. It requires a certain supply of unrestricted funds. A group of skilled persons compete in trying to develop a formula which most successfully explains the configuration and structure of a hand. There is a pool of genes from which one can draw or they are distributed, and the contestant attempts to fit them into sequences or combinations. They have different degrees of penetrance. An effort is made to establish linkages, sometimes known as pairs; they may even be recognized by the theorist in peculiar positions—for example there is a pairing known as back to back which has a special power. But in addition, the effort may be to find the correct gene to fit a given locus.

In some sequences the dominant characteristic results in a hand with a flush; in others the proper arrangement emphasizes the dominant feature of the hand as straight. In these situations an effort is made to match the items as homozygotes, and this adds further power whereas the appearance of a heterozygote may not provide the essential combination and the experimenter, recognizing that his hypothesis of hand formation is not further tenable, will announce that he drops out and forfeits his chance and will shortly undertake a new formulation.

Another type of combination is seen when the same code number may reappear in successions and pairs. This may result in an effort toward explanation known as Queens full of genes.

There is a distinct statistical probablility which applies to these formulations, but at times the concept of wild genes is introduced, meaning that certain of them can be accepted in any locus and with increased dominance, so that strong combinations may rapidly be formed and their value supported, but the statistical significance is often difficult to calculate.

The investigator who completes and defends his formula of hand formation most effectively is rewarded by a modified Mendelian prize—not of colored beans, but a collection of discs. Instead of black and white, however, they are commonly blue, red, and white. The successful contestant then arranges them in piles and calculates their distribution until the next exercise is begun. As a result of his successful manipulations of the genes, the winner earns the title, Genius.

This brief and incomplete discussion can only mention that the exercise has different sets of guidelines. One form is called penny ante—named for the character in a child's book, "Henny Penny." It means the contestants are chicken. Another level which is undertaken with grave demeanor and full appreciation of the values involved commonly requires limited light, a haze of tobacco incense and sometimes other stimuli, so that the persons will poke about in the gene pool with better appreciation of its content. I could describe this better had I a better knowledge of the genetic code.

This is to welcome you to the game and hope that you will defend your formulations with judgment and enthusiasm, recognizing, however, that a sequence of four does not gain acceptance at full value. Other contestants may demand full evidence or offer competing information of greater power. Ladies and gentlemen, I shall now proceed to open.

We have previously described epidemiology as the study of the laws and factors governing the occurrence and distribution of disease and disorder in a population. These factors include the characteristics of the population, the causative agencies, and the biological, social, and physical environment. Of all the factors which may be involved in the occurrence or absence of disease, the genetic influence is probably the most pervasive. Our very presence here to discuss the effect of

our past upon our future is the consequence of our heritage. But even casual reflection leads to the realization that this survival cannot be attributed to any one item in our genetic code; rather the influences are multiple or polygenic. It is the combination which counts. This adds to the complexity of an already complicated problem for, as Mayr (1961) points out, there is as yet no adequate analysis of a single polygenic trait in man, although they include characteristics that are particularly important to mankind. Similarly, there are few, if any, disorders whose occurrence is *not* the resultant of multiple convergent factors. So when the human geneticist turns to disease and disorder in the population as his basis of genetic analysis, he is promptly in epidemiology. And when he asserts the concept of multiple factors to produce an effect, he is in full cry epidemiologically. Conversely, where the epidemiologist seeks explanation for familial or other group aggregations of health or disease, he is immediately involved in genetic problems. In each case the investigator is concerned with "population thinking" (Mayr, 1961), and the beginning of formation of an epidemosis (Francis, 1959).

An inherited characteristic may be brought to view only by appropriate challenge. It has been repeatedly emphasized that whether a gene is harmful or useful or indifferent is commonly related to the environment in which the bearer is exposed. The genetic epidemiologist then becomes ecologist seeking significant correlations between a group disorder and one or other stress from the great array of environmental influences. The success of these efforts will obviously be related to the uniqueness of the stress and the directness of its effect, to the frequency of the basic defect and to the ease of detection of the disorder under consideration. Geographic variations may be valuable but also confusing. And even if one finds a correlation, perhaps all he can say is, "Redheads are more sensitive to heat." It is clear that this type of extensive and intensive investigation is not entered into readily, unless the problem is of major importance. Nevertheless, it is the starting point for mapping the numerous genetic factors which go into determining resistance or susceptibility of a population to pathogenic stresses in the environment. In a heterogeneous population and a rapidly changing environment the possibilities seem infinite.

The difficulty in establishing genetic correlations for certain disorders, as distinct from environmental effects, has been well recognized. Nevertheless, it is one of the basic hopes of the Community Health Study which we are conducting in Tecumseh, Michigan. Taking atherosclerosis and coronary heart disease as an example, the hypothesis is briefly this: Atherosclerosis and coronary heart disease are the result of a basic metabolic disturbance. Despite much discussion and inferential data, it is not established that coronary heart disease has a significant familial distribution nor has the weight of genetic factors

versus environmental influences in causation been adequately determined. Nevertheless, when a case of coronary heart disease occurs, the probability is that additional cases are more likely to occur in the family of the index case than in similar families in which there is no identified case. This should be true whether the determining factor is genetic or environmental, since the household is more likely to be sharing a common environment and way of life; the blood kindred is also more likely to be sharing the same genetic determinants. The kindred thus constitutes a pool of susceptibles for continued study in comparison with other unaffected kindreds.

The total community was identified as individuals, as households, and as genetic kindreds. In 1959–60 extensive histories were obtained and careful physical examinations were made by physicians from the clinical departments of the medical school in the clinic established at Tecumseh. A variety of laboratory tests included X-ray, EKG, determinations of blood cholesterol, sugar, uric acid, proteins, and urine examinations. In this community of 10,000 people there are 4,500 kindreds and something over 2,000 households. Nearly 90 percent of all ages participated in the study, and those who didn't were fully identified. Diagnostic criteria were carefully established and utilized after complete review of the findings. The entire procedure has required great effort and much time, so that many of the fascinating data are but now becoming available for proper analyses. One feature of value has been the division of the population into ten representative samples, so that the rotation through the clinic always constitutes a representative segment of the population. Special studies can then be conducted without having to draw new samples and without involving the entire population.

The second round of interviews and examinations began a year ago. They have been enlarged in certain details of the history and in some phases of the physical and laboratory examinations. The participation continues high. A continued surveillance of the community and the people is also being established.

There is need to emphasize that the components of special interest come out of this natural community in a natural unselected manner. They exist in a natural ecologic setting. Their position as individuals, as households or kindreds in the study derives from the data obtained by methods of interview and examination uniformly applied to the total population, not by prior selection. The various data can be viewed independent of any clinical diagnosis, as indices of suspicion, by observing the subsequent course of those with and without the particular characteristic. It is here that the oncoming members, the children and younger segments, of the hypothetical susceptibles can be further compared with matched groups or kindreds of presumably nonsusceptibles. In this manner the natural risk of developing a disorder can be measured against a general ecological background

which will hopefully bring out the individual or group variation. It is expected that the role of the genetic factor in aggregation of disorder can be established not only for coronary heart disease, but for diabetes, hypertension, chronic pulmonary disease, rheumatoid disease, and others. They have numerous overlapping features, so that etiologic relationships of various abnormalities and their combinations may be considered in a common host population. For example, are coronary heart disease and diabetes different manifestations of the same basic disorder, of the same polymorphism?

A still further objective is to find new ways of determining the nature of susceptibility through research in basic science—physiologic, biochemical, physical, etc., applied to the community; and as already indicated, efforts will be made to detect the beginnings or precursory stages. Here the careful continued study of the younger age groups and the respective cohorts of the kindreds are expected to be the most rewarding in the search for, or application of, preventive measures. Continued ecologic observation should also aid in revealing natural exposures or changes which would constitute environmental pathogenic agencies. In support of that objective a mapping of the biological, social, and physical environment is being made. The role of physical activity, as well as the social and psychological reactions of the individuals are to be incorporated with other data of possible importance. I think it is safe to say that comprehensive epidemiological information of high clinical excellence is for the first time becoming available from a total natural community.

The question of numbers is always before us, but our purpose is not merely one of measuring incidence. It is fair to say that there are enough cases of our target disorders to provide a good collection of potential susceptibles for detailed and continued study. Attention to these and their comparison groups will be enhanced in the sense of biological and clinical investigation, while that to the total population can be reduced in intensity. The prospects for the ongoing study are that the genetic information from this epidemiological base will contribute new, valuable data for comparison with environmental influences at the same time.

It is apparent, however, that the pathologic index of coronary occlusion or even probable coronary heart disease may be unsuitable for analysis of oligogenic determinants. The case of acute coronary occlusion may be, and I suspect this is so, only the accident in the course of a basic aberration present in the population generally. In that event it would require that the genetic determinant apply to the accident rather than to the underlying disturbance, which would seem more difficult of achievement. This emphasizes also that for our purposes it is necessary to project inquiry beyond a demonstration of a correlation between some factor and a result, in order to determine

the pathogenic mechanism which may then be approached prophylactically or therapeutically.

The influence of environment upon genetic expression is generally considered in terms of the external environment. The genetic makeup and potential, on the other hand, is thought of as part of the internal environment, intrinsic, inherent, and immutable. This ignores the possibility that other components of the internal environment may play a role in altering the genetic effect; that mutations may not be quite as chancy as they seem. It is not within my competence to discuss this in detail except to point out that a variety of presumably external influences may have an intimate relation to the internal genetic mechanisms. Increasing knowledge of viral infections and cellular biology demonstrate the manner in which viruses may participate in and alter genetic mechanisms of cells, particularly when they persist in a temperate or latent stage as part of the internal or intracellular environment. This may result in new karyotypes, in new antigens, in escape from general regulatory mechanisms. Viruses may exert genic repressions. At what stage do these somatic effects influence organismal heredity? The dominant influence genetics had earlier exerted upon the interpretation of spontaneous tumors or leukemia in different strains of mice may in reality relate more to susceptibility to a virus than to tumor production *per se.* Moreover, viruses may be constant accompaniments of the reproductive process. Can a virus which has wide distribution through the blood stream induce sufficient variation in the germ cells to result in transmissible changes in the code? I retain a rather lurking suspicion that the malarial parasite has a role in inducing the sickle-cell change rather than only as a selective phenomenon. Chemical and physical agents raise the same questions. But this is only a brief inquiry into problems of genetic epidemiology where the possibility of induced mutation under natural conditions rather than evolutionary selection of the spontaneous mutant is considered. It emphasizes, however, some of the additional considerations which arise when one attempts to study all the factors which may govern the occurrence and distribution of a genetic disorder in a population.

BIBLIOGRAPHY

1. Francis, T., Jr., 1959, The Epidemiological Approach to Human Ecology, Am. J. Med. Sci., *237:* 677–684.
2. Mayr, E., 1961, Evaluation and Man's Progress, Daedalus, Proc. Am. Acad. Arts and Sci., *90:* 460–461.

Genetic Studies of Family Units

By Bertram L. Hanna,* Ph. D., *Departments of Medicine and Neurology, Indiana University, Indianapolis, Ind.*

The common interest of the participants in this symposium is the etiology of chronic disease. The factors responsible for the production of an observed state in the individual may be broadly classified as (1) those which are intrinsic, transmitted to the zygote from which he developed through a marvelous series of steps which, though thoroughly studied during the past century are not yet fully comprehended, and (2) the myriad of extrinsic determiners, known and unknown, which have acted from conception to the time of examination to modify the organization of the individual. We wish today to devote our discussion to the intrinsic determiners of disease, realizing full well that it is impossible in many cases to ignore all of those environmental variables and stresses continually operating upon the human organism.

GENERAL CONSIDERATIONS

The epidemiologic approach to the study of etiology consists essentially of comparing the distributions of disease in groups characterized by similarities of age, sex, race, location, occupation, or ethnic background in an attempt to characterize diseased and nondiseased groups and to discern those factors which may be correlated with the occurrence of disease in a population. The occurrence of family

*Dr. Hanna died in February 1964 of complications of diabetes. This is one of his final scientific contributions. His death is keenly felt by all of his colleagues, but perhaps particularly so by the participants in this symposium, who will recall vividly the quiet courage with which he carried on at this meeting despite an attack of acute glaucoma.

aggregations in the population provides information, generally overlooked in epidemiological studies, which will permit the examination of inherited intrinsic determiners which are of significance in the etiology of disease. The binomial nature of the genetic system in sexually reproducing organisms permits the construction of simple mathematical models for the genotypic expectation among the offspring of matings or in a population. The method of genetic analysis thus consists of a comparison of sample statistics with parameters specified by genetic theory. In the strictest sense the genetic study of family units includes only studies of segregation and linkage. In practice, however, the geneticist frequently utilizes the armamentarium of the epidemiologist to obtain ancillary information both from families and from the population which enables him, by assuming specific genetic models, to examine various problems not directly amenable to the more precise methodologies.

Several authors have listed criteria suggestive of genetic factors in traits of unknown etiology (David and Snyder, 1952; Neel and Schull, 1954; Steinberg, 1959; Morton, 1962). Among these are:

(1) A greater frequency of occurrence in relatives than in the general population.
(2) Greater concordance in identical than in fraternal twins.
(3) Increased frequency of defects at the same site or of the same functional system in relatives as compared with the general population.
(4) An excess of consanguineous parentage for a presumed recessive trait.
(5) Onset at a characteristic age without a known precipitating event.

It is obvious that no one of these relationships provides, either singly or in any combination, the critical evidence required to establish a genetic etiology for disease. Despite every effort to rule out environmental determination there will always remain in the mind of the cautious investigator some doubt of the validity of any conclusion based upon such criteria. These "epidemiologic impressions" should, however, suggest more detailed examination of family patterns.

Evidence for a genetic etiology may be obtained by demonstrating that (1) the observed distribution of offspring in families is compatible with the expectation under a specific genetic model, or (2) the distribution of phenotypes in the population is compatible with the distribution expected under the same genetic model. A good fit of the observed data with a genetic model does not conclusively demonstrate a genetic etiology. The genetic approach is, however, quite powerful since the nonrandom genetic variation is incorporated in the structure of the model, with environmental variation being considered random among family members. In unsophisticated epidemiologic studies, genetic variation is assumed to be random;

the power of such studies could be greatly increased by incorporating into their design knowledge of family relationships derived from genetic theory (Fisher, 1918; Kempthorne, 1957).

SEGREGATION MODELS AND MODES OF INHERITANCE

The segregation models employed for the analysis of traits produced by genes not located on the sex chromosome are summarized in table 1, in which two alternative types of genes (alleles) at a single locus are designated *A* and *a*. Matings in the population fall into six different classes as listed under the column headed "parental genotypes" in table 1. The expectation for offspring differs for each of these mating types. If it is possible to identify the genotypes (*AA*, *Aa*, and *aa*), of parents and of all individuals within sibships, then there is little difficulty in testing the hypothesis that traits *A* and *a* result from two alleles at the same locus which segregate according to genetic expectation within families. An investigator simply examines his families, assigns them to the proper mating type groups and then does a simple statistical test of goodness of fit to determine whether the frequencies observed among the offspring are compatible with expectation under the model.

TABLE 1.—*Models for the genetic analysis of autosomal traits with family data*

Parental genotypes	Offspring expectation		
	AA	Aa	aa
AA × AA	1	—	—
AA × Aa	1/2	1/2	—
Aa × Aa	1/4	1/2	1/4
AA × aa	—	1	—
Aa × aa	—	1/2	1/2
aa × aa	—	—	1

In practice, the use of segregation models is rarely so simple and straightforward. The complicating factors may be classified broadly as (1) dominance, (2) ascertainment, (3) interactions, and (4) multiple etiologies. While each demonstrable allele at a segregating locus must have some effect upon the structure or function of the organism, the activity differential within an allelic series need not be straightforward. Suppose, for example, that an allele *A* results in the production of two units of demonstrable cellular protein, its allele *a* producing one. If the dominance relationship at this locus is additive, then assays ranging from two (genotype *aa*) through four (genotype *AA*) are expected in the population. A deviation from additive dominance might occur, however, such that genotype *Aa* has a value of four, suggesting that the allele *a* acts differently in the presence of *A* than when paired with another *a*.

735-712 O—65——2

In addition to true "gene dominance," we may speak of trait dominance. This will appear for traits for which the proper tools for heterozygote recognition are not available to the investigator. For our discussion today we will speak of dominant and recessive traits rather than dominant and recessive genes since the trait is the observed event from which the investigation of etiology proceeds.

The problems of ascertainment will be discussed in detail by Dr. Crow later in this program. We shall only note here that family sampling is complicated by dominance relationships, by the small size of human families, which permit some parents potentially capable of producing affected offspring to be omitted from a sample, and by differing probabilities of including in a sample families having more than one affected offspring. Family samples collected under various types of ascertainment must thus be treated statistically in different manners to obtain comparable reliable results.

The investigator is generally unable to measure directly the action of a single gene locus. He is required to work with end-products, produced through multiple developmental and physiologic interactions during the lifetimes of the individuals. Environmental factors or other genes segregating in a family may produce variation in the expression of a genetic trait or may prevent the appearance of symptoms in a genetically susceptible individual (incomplete penetrance). Variable expression presents no problem for genetic analysis if the different manifestations are recognizable, but incomplete penetrance, since it leads to a modification of the observed ratio among the offspring, is a serious problem. Genetic analysis is impossible if penetrance is low.

If a disease having delayed onset is rare, the penetrance may be calculated from the distribution of age of onset in family members, without any genetic assumptions (Morton, 1957). If incomplete penetrance is entirely due to delayed onset, the average penetrance lies between $\int f(z)G(z)dz$ and $\int f_1(z)G(z)dz$ where $f(z)$ is the frequency of age z at death or last examination among normal and affected siblings, $f_1(z)$ is the same frequency with the index cases excluded, and $G(z)$ is the cumulative frequency of onset at age z among affected cases. Heterogeneity in age of onset among families will cause such a penetrance estimate to be approximate only (Morton, 1957). For a dominant trait which occasionally skips a generation, penetrance may be estimated from the number of unaffected carriers in the direct line of descent from an affected ancestor to the first affected descendant in that line, if the trait is sufficiently rare that the chance of a "normal" carrier mating into the family is negligible. If the penetrance is low, however, long skips may fail to be recorded so that the penetrance will be overestimated.

A proportion of the cases observed in a body of data may have arisen through nonrandom nonrecurrent events such as mutation,

chance expression of a heterozygote, or purely environmental effects such as prenatal anoxia, maternal-fetal incompatibility, toxic drug therapy, or trauma. With respect to family analysis these are all "true phenocopies" and should, where possible, be excluded from the analysis.

Families examined for a test of genetic etiology may be divided broadly into two groups. Some sibships will include two or more affected children and have been called multiplex families by Morton (1959). Other families will contain only a single affected child (simplex families) which may occur as a result of a number of different precipitating events. A certain proportion of them will have the same genetic etiology as do those cases in the high risk families, but others may occur as a result of environmental factors, illegitimacy, errors in diagnosis, or from genetic causes other than those found in the high risk families. Morton distinguishes these two groups as chance isolated and sporadic families. Methods are available for estimating the proportion of sporadic cases in the data, making use of the distribution of affected children in multiplex and simplex families (Morton, 1959). These methods will be discussed by Dr. Crow in the following paper.

For many diseases of clinical interest the proper diagnostic tools are not yet available to distinguish between two independently inherited diseases leading to the same observable abnormality. If one of these is autosomal and the other sex linked, genetic separation is not difficult, but two diseases inherited in the same manner will defy genetic analysis unless both are rare, in which case the chance of both occurring in one sibship is remote.

AUTOSOMAL DOMINANT INHERITANCE

The inability of the investigator to recognize a heterozygote requires modification of the model in order to test genetic hypotheses. If we now recognize *A* in table 1 as a dominant gene, it will be seen that for an autosomal dominant trait, affected children will occur only when at least one parent is affected. Thus for an autosomal dominant trait, transmission is directly from one generation to the next with the two sexes equally affected and with each affected child having an affected parent except for the rare case in which mutation occurs. Because of the limitation on human family size, not all families with an affected parent will be expected to have an affected child. If the trait is sufficiently rare in the population it may be further assumed that every affected individual is heterozygous. If this is the case the majority of families observed will be of parental genotypes $Dd \times dd$ and will have a segregation frequency of one half. This assumption appears reasonable except for dominant traits

which are quite widespread in the population. It may be shown, for example, that if the incidence of an autosomal dominant condition in the population is 1/5,000 individuals, the probability of including in a sample of one hundred families any marriages other than those between heterozygous affected individuals and homozygous normals is less than one in twenty. For conditions having a frequency of less than 1/50,000 the risk for an irregular sample of any reasonable size is negligible.

The model for the test of genetic segregation for a rare dominant trait is a true genetic model, since the offspring expectations are derived directly from genetic theory. A simple statistical test of significance is sufficient to demonstrate whether or not one may accept the null hypothesis that the trait in question is inherited in this manner. If a dominant trait is common in the population so that the frequency of affected-by-affected matings is elevated, it is necessary to obtain an estimate of the frequency of the condition in the population of marriageable age in order to make tests of genetic hypotheses. In the absence of evidence of assortative mating the Hardy-Weinberg equilibrium model may be employed to estimate allelic frequencies for the locus in question. From these the expectation in the sample of families may be obtained as shown in table 2. Because of limited family size the distribution of offspring in a family will not necessarily identify the parental genotype, so that in a sample of families only two mating types may be distinguished; affected $\times$ affected and affected $\times$ normal. This method lacks the power of the one previously discussed since an estimate of a population parameter, which itself has sampling variance, is taken as the basis for model construction. An analysis of this type should be supported by a population analysis to demonstrate that the observed frequencies are distributed as expected under the proposed genetic model. The early acceptance of the genetics of the ABO blood groups was based in large measure on parallel family and population analyses of this type (Bernstein, 1925; Strandskov, 1931).

AUTOSOMAL RECESSIVE INHERITANCE

A condition which is recognized as an autosomal recessive trait will occur among children from three types of matings in the population (table 1). If the recessive genes are allelic, all children of affected parents will be affected, regardless of the sex of the children. If both parents are normal heterozygotes, the proportion of affected offspring is 1/4; if one parent is affected, it is 1/2. If the trait is rare, most of the parents and relatives of an affected sibship will be normal. For rare recessive diseases the rate of parental consanguinity is elevated. A carrier of a rare gene is unlikely to marry a carrier of

TABLE 2.—*Models for the genetic analysis of common autosomal dominant traits with family data, where q^2=the frequency of unaffected individuals in the population, and $p+q=1$*

Parental matings		Frequency of matings	Offspring expectation		
Phenotypes	Genotypes		AA	Aa	aa
Dominant × Dominant	AA × AA	p^4	p^4		
	AA × Aa	$4p^3q$	$2p^3q$	$2p^3q$	
	Aa × Aa	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
			p^2	$p^2q(1+p)$	p^2q^2
Dominant × Recessive	AA × aa	$2p^2q^2$		$2p^2q^2$	
	Aa × aa	$4pq^3$		$2pq^3$	$2pq^3$
				$2pq^2$	$2pq^3$
Recessive × Recessive	aa × aa	q^4			q^4

EXPECTATION IN FAMILIES

	Affected	Normal
(a) One parent affected	$1/(1+q)$	$q/(1+q)$
(b) Both parents affected	$(1+2q)/(1+q)^2$	$q^2/(1+q)^2$

the same gene if he mates at random; the likelihood is greatly increased if he marries a relative. It can be shown that the proportion of cases of a recessive disease in the population which result from consanguineous marriage is approximately $\alpha/(q+\alpha)$, where q is the frequency of the gene in the population and α is the population inbreeding coefficient. Since $\lim \alpha/(q+\alpha)=1$ as q approaches zero, the proportion of cases derived from consanguineous marriage provides evidence for the relative incidence of the condition in the population.

It will be seen from table 1 that children with a recessive disorder can result from only three types of parental matings. These matings may be separated for analysis by examination of the parental genotypes. The mating of an affected by a homozygous normal parent will not be included in a body of data ascertained through the children, but will constitute the larger proportion of affected by normal matings obtained in data ascertained through the parents, particularly if the condition is rare in the population. For a rare recessive trait, the majority of matings producing abnormal children will be between phenotypically normal heterozygotes. If ascertainment is through the children, families with only one child must be excluded since the fact that every such family has an affected child will lead to a serious overestimate of the segregation ratio. If, on the other hand, ascertainment is through the parents, such families need not be excluded.

Methods for the family analysis of recessive traits are derived entirely from genetic theory and are therefore quite powerful. These methodologies have been reviewed by Neel and Schull (1954), Li (1961), Steinberg (1959), and Morton (1959, 1962) and will not be discussed further here.

Models employed for the analysis of traits produced by genes located on the *X* chromosome are summarized in table 3; segregating alleles may be designated X^A and X^a to indicate dominant and recessive alleles located on the *X* chromosome. It will be seen from this table that a male offspring always receives his *Y* chromosome from his father and his *X* chromosome from his mother. Each female offspring carries the *X* chromosome of her father and has a 1/2 chance of carrying either *X* chromosome of her mother. In a population, the distribution of *X* chromosomes in males is thus a random sample of the *X* chromosomes of their mothers, presumably randomly selected from among the females of the previous generation. If the trait is recessive and sex linked, most cases will occur in the male offspring of normal parents, and no relatives will be affected except males on the maternal side of the family. Affected males will not transmit the trait to either sex but all of their daughters will be carriers. The frequency of a sex-linked recessive trait in females will be approximately the square of the frequency in males. Excluding *XO* females (Turner's syndrome), affected females will arise only from matings of affected fathers and carrier mothers and, if the trait is rare, there will be elevated consanguinity in families with affected females.

If a condition is inherited as a sex-linked dominant trait, every affected offspring will have an affected parent, barring mutation and misconception. If the condition is rare, the frequency in females will be approximately twice the frequency in males in the population. If a mother is affected, the expectation is that 1/2 of her children of both sexes will be affected, but affected fathers will transmit the condition to all of their daughters and none of their sons.

TABLE 3.—*Models for the genetic analysis of sex-linked traits with family data*

Parental genotypes		Offspring expectation	
Mother	Father	Female	Male
X^AX^A	$X^A(Y)$	X^AX^A	$X^A(Y)$
X^AX^a	$X^A(Y)$	$1/2X^AX^A$ $1/2X^AX^a$	$1/2X^A(Y)$ $1/2X^a(Y)$
X^AX^A	$X^a(Y)$	X^AX^a	$X^A(Y)$
X^AX^a	$X^a(Y)$	$1/2X^AX^a$ $1/2X^aX^a$	$1/2X^A(Y)$ $1/2X^a(Y)$
X^aX^a	$X^A(Y)$	X^AX^a	$X^a(Y)$
X^aX^a	$X^a(Y)$	X^aX^a	$X^a(Y)$

Other types of sex-linked inheritance have been suggested to explain the patterns observed for various conditions. In the early literature, several pedigrees appeared which suggested the occurrence of a condition resulting from a gene located on the *Y* chromosome (*Y* linkage). Such a gene would be transmitted directly from a father to all of his sons and could never occur in a female. Stern (1957)

reviewed this early work and concluded that there is no evidence to support the claim for *Y* linkage for most of the conditions examined. Only one condition, the occurrence of hairy pinnae has been reexamined in recent years, but critical evidence in support of *Y* linkage has not yet been forthcoming. Haldane (1936) suggested that there might be occasional crossing over between homologous portions of the *X* and *Y* chromosome in man, producing partial sex-linked inheritance patterns. Critical reevaluation of earlier studies by Morton (1957a) casts serious doubt upon the validity of this hypothesis.

SEX LIMITED INHERITANCE

Some autosomal traits are limited in their expression to one sex. These are generally associated in some manner with sexual differentiation or reproductive physiology. If the trait is autosomal recessive, all the criteria of this mode of inheritance apply except that only one sex is at risk. If the trait is autosomal dominant and is limited to males, one half of the sons of unaffected carrier females will be affected so that a pedigree may, on cursory examination, suggest sex-linked inheritance. One half of the sons of an affected father would, however, be affected for an autosomal dominant trait, whereas all children would be normal were the trait sex-linked recessive. In addition, an affected *XO* daughter could occur through nondisjunction in a carrier for a sex-linked recessive but could not occur if the trait is sex limited.

LINKAGE MODELS

The association of inherited traits in members of a family affords an opportunity to investigate genetic linkage, which is the occurrence of genes responsible for the traits at loci near each other on the same chromosome. It is not specific alleles which are linked, however, but genic loci. The association of a disease condition with a specific blood type in a family provides no evidence for linkage since other blood types resulting from alleles will also segregate with the disease if the responsible loci are on the same chromosome. A test of the null hypothesis of genetic independence may, however, be made with families in which at least one parent segregates both traits. The expectation of offspring from the cross of double heterozygote by homozygote parents for linked loci having a crossover frequency *c*, is given in table 4. From data of this type an estimate of *c* may be obtained directly.

The test of linkage in humans is complicated by the fact that it is only rarely known whether a parent is of genotype *AB/ab* (coupling

TABLE 4.—*Models for the analysis of linkage with backcross matings of (a) coupling and (b) repulsion phase, where c is the frequency of recombination between two loci on the same chromosome*

Parental genotypes	Offspring expectation			
	AB/ab	Ab/ab	aB/ab	ab/ab
(a) AB/ab × ab/ab	(1−c)/2	c/2	c/2	(1−c)/2
(b) Ab/aB × ab/aB	c/2	(1−c)/2	(1−c)/2	c/2

phase) or *Ab/aB* (repulsion phase). In addition, human families are rarely large enough to permit the total genetic potentiality of the parents to be expressed. Bernstein (1931) devised the earliest method of scoring families independent of phase to provide a test of the null hypothesis. The historical development of linkage analysis in humans has been reviewed by Morton (1955) who developed a sequential probability ratio test requiring a much smaller sample size for a given risk of error than do earlier methods (Morton 1955, 1956, 1957*a*). Smith (1959) has applied Bayes' theorem to obtain, with Morton's tables of likelihood scores, an estimate of probability for the occurrence of linkage.

Only two autosomal linkages have been substantiated to date. Renwick and Lawler (1955) demonstrated linkage between the loci for nail-patella syndrome and the ABO blood groups, and Morton (1956) showed linkage between the loci for elliptocytosis and the Rh blood groups. Elliptocytosis was shown to be very closely linked to Rh in four large pedigrees but to be independent in three others, suggesting that at least two different genic loci may be responsible for this condition.

INCIDENCE OF GENETIC DISEASE

The incidence of disease has always been of major interest in the study of epidemiology. The use of family models for the study of diseases of known genetic etiology may greatly increase the precision of such estimates. Studies of the frequency of cystic fibrosis, a chronic and usually fatal disease of children have produced prevalence estimates of from 1/10,000 (Selander, 1962) to 1/100 individuals (Goodman and Reed, 1952). One excellent epidemiologic study suggests an incidence of approximately 1/3,700 live births (Steinberg and Brown, 1960). These authors summarize succinctly the deficiencies of their method, noting the absence of confirmed diagnoses, the incompleteness of sampling, the problems of migration, and the inability to determine exactly the initial population represented by their affected group. Our studies of cystic fibrosis further point up the problems faced by the epidemiologist. In a carefully controlled study of cystic fibrosis

sibships carried on during the past 2 years, approximately 44 percent of 170 probands referred to us by local physicians, hospitals, and National Cystic Fibrosis Research Foundation chapters have been found, on careful reexamination with sweat tests confirmed by independent laboratories, to have something other than cystic fibrosis despite clinical findings which on the surface might suggest this diagnosis (Merritt *et al.*, 1963). It is apparent that this diagnosis has become popular in the past few years, although it may not have been so at the time of earlier studies.

Cystic fibrosis appears to segregate in families as an autosomal recessive trait (Anderson and Hodges, 1946; Lowe *et al.*, 1949). The absence of a significant increase in parental consanguinity suggests that the gene may be relatively common in the population. There is general agreement among the workers in this country that despite improvements in clinical management in recent years the likelihood is quite small that an affected individual has survived to reproduce. The evidence from our data suggests that the selection against affected individuals prior to the time of diagnosis is about 1.6 times as great as that of the normal individual. The mutation rate of 1×10^{-3} estimated by Goodman and Reed (1952) is inordinately high and we believe, together with these authors and with Steinberg and Brown (1960), that the concept of heterozygote advantage is perhaps more realistic at the present time. If mutation at this locus is of the magnitude reliably reported for other loci in man, selection of less than 2 percent against the normal homozygote is sufficient to maintain cystic fibrosis in the population. The inherited nature of cystic fibrosis is not widely known to the public and no evidence of assortative mating is seen in the present generation.

Each parent of a cystic proband must be a carrier of the responsible gene and, barring mutation, must have received this from a carrier grandparent. Each of the siblings of this parent will therefore have a precise probability of being a carrier which is specified by genetic theory. Examination of the frequency of cystic fibrosis among the offspring of these siblings therefore gives an estimate of the frequency of cystic fibrosis carriers in the population and of the frequency of the mutant allele producing this disease.

Two different methods are available for this analysis. These may be called (1) the population or single selection method, and (2) the segregation or complete selection method. These differ in the assumptions required for model construction as given in table 5. Let P be the proportion of homozygous normal and Q the proportion of heterozygous carrier individuals in the population of marriageable age. For the population model, P and Q are constant in successive generations. The probability that a carrier individual mates with a normal homozygote is P; with an Aa carrier, Q. The probability of being an Aa

carrier is then $\frac{2}{4-Q}$ for each of the siblings of a known carrier parent. The risk that a sibling will marry a carrier in the population is Q, so that the probability of a first cousin of the proband being affected is $2/(4-Q)\times Q\times\frac{1}{4}=\frac{Q}{2(4-Q)}$. Since $Q/2=q$ and $(2P+Q)/2=1-q$ where q is the frequency for the mutant allele, the risk for a first cousin is $\frac{q}{2(2-q)}$. From this we estimate $q=4R/(T+2R)$ where R is the number affected among T cousins in a random sample of cousin sibships.

TABLE 5.—*Assumptions regarding the population and the families required for the estimation of allelic frequency from first cousin data by alternative methods*

Assumption about the—	Population method	Segregation method
A. Population	1. Random mating with respect to trait 2. Genetic equilibrium	1. Random mating with respect to trait.
B. Family	1. Negligible mutation 2. Uniform parental fertility 3. Random survival of parents to reproductive age.	1. Negligible mutation. 2. Uniform parental fertility.

This model assumes genetic equilibrium in the population, but no such assumption is made for families; the information required to permit correction for mutation or differential selection within families is not available at present. In addition, the required random sample of offspring sibships from the parental generation is difficult and expensive to obtain. The reliability of estimates obtained by this method may thus be questioned, even if extremely large samples are examined.

For the segregation model, a sibling of a known carrier parent is taken as proband for the sibship consisting of his children, which is thus ascertained independent of the occurrence of cystic fibrosis. If it is assumed that all grandparental matings are $Aa\times AA$, an estimate of the maximum possible allelic frequency in the population is obtained. If it is assumed that all grandparental matings are $Aa\times Aa$, a minimum estimate is obtained. These estimates and their standard errors are given in table 6. It is clear that the values derived from the segregation model provide unbiased estimates of the maximum and minimum allelic frequencies possible in the parental generation of marriageable age. The "correct" allelic frequency (q^*) lies between these two; as q^* decreases in magnitude the difference q^*-q_{max} approaches zero. For frequencies in the range found for cystic fibrosis, the difference between q_{max} and the frequency estimated from data meeting all conditions specified for the population model is approximately 5×10^{-4}.

TABLE 6.—*The estimation of maximum* ($\overline{q}$) *and minimum* ($\underline{q}$) *frequencies of the allele for cystic fibrosis from cousin data. R is the number of affected among T cousins examined*

Assumed grandparental matings	Allelic frequency	Variance
AA × Aa (maximum estimate)	$\overline{q} = 4R/T$	$V(\overline{q}) = \overline{q}(4-\overline{q})/T$
Aa × Aa (minimum estimate)	$\underline{q} = 3R/T$	$V(\underline{q}) = \underline{q}(3-\underline{q})/T$

The major advantage of the segregation method is that if the occurrence or nonoccurrence of affected cousins is not a criterion for the inclusion of sibships in the body of data, random sampling of the parental generation is not required. No assumptions regarding the population equilibrium or selection within families are required unless one attempts to interpolate a "corrected" frequency from the estimated maximum and minimum values.

The incidence among conceptions in the offspring generation is then $I=q^2$ with 95 percent confidence limits of $(q\pm\sigma q)^2$. Estimates obtained from a sample of 982 cousins of 74 cystic fibrosis proband sibships are given in table 7. These data indicate an incidence of 15/100,000; the probability that the incidence of cystic fibrosis in the population represented by this sample is greater than 69/100,000 is less than 0.025.

TABLE 7.—*The maximum incidence of cystic fibrosis estimated from the frequency of occurrence among first cousins of affected probands*

	Sibships	Total children	Affected children
Proband	74	233	95
1st cousin	332	T=982	R=3

q max=12/982=.0122±0.014.
I∼.00015 with 95 percent upper confidence limit∼.00069.

CONCLUSIONS

The familial clustering of disease may result from genetic or environmental factors or from complex interaction of intrinsic and extrinsic factors. Although an increased frequency of occurrence in the relatives of an affected individual is suggestive, detailed analysis of carefully collected family data is required to provide a critical test of genetic etiology. Examination of the distribution of phenotypes within sibships or in the population will frequently suggest which genetic model may be appropriate. Simple statistical methods are frequently adequate to test the goodness of fit of sample statistics to parameters specified by genetic theory.

The analysis of family data is generally complicated by the inability of the investigator to specify the genotypes of individuals, by devia-

tions in sampling procedure and by the interaction of environmental factors or of other genetic factors segregating in families. Powerful statistical methods provide a means of circumventing the first two of these problems with carefully collected data but interactions, either environmental or genetic, may make analysis impossible with currently available methods. The segregation methods discussed here are not suitable for the analysis of many of the "familial" chronic diseases, which are known to be affected to a large degree by environmental factors but which may, in addition, be genetically complex. The number of qualified investigators now interested in the etiology of these diseases would suggest that problems such as the diagnosis of preclinical or latent states and carrier diagnosis may, in time, be solved for many of them so that meaningful segregation and linkage studies may be done.

The human geneticist, in his study of families and populations, frequently relies heavily upon well-established epidemiologic procedures. Knowledge of genetic theory and of the methodology of human genetics may likewise contribute to epidemiology by suggesting studies in which the nonrandom intrinsic variation resulting from family grouping in the population may be either partitioned out or isolated for detailed study. Both the geneticist and the epidemiologist may, however, be unaware of many of the pitfalls with which the other, by training and experience, is thoroughly familiar. It is clear that many problems in the etiology of the chronic diseases require close cooperation between the two. It is to be hoped that significant advances in our knowledge may result from such cooperation.

REFERENCES

1. Anderson, D. H. and Hodges, R. G., 1946. Celiac syndrome. V. Genetics of cystic fibrosis of the pancreas with a consideration of etiology. *Amer. J. Dis. Child.* 72: 62–80.
2. Bernstein, F., 1925. Susammenfassende Beotrachtungen über die erblichen Blutstrukturen des Menschen. *Ztschr. indukt. abstamm. u. Vererb.* 37: 237–270.
3. Bernstein, F., 1931. Zur Grundlegung der Chromosomentheorie der Vererbung beim Menschen mit besondere Berucksichtung der Blutgruppen. *Z. indukt. abstamm. u. Vererb.* 57: 113–138.
4. David, P. R. and Snyder, L. S., 1952. Genetics and disease. Proc. Second Canc. Cong., Am. Canc. Soc. N.Y. 1128–1138.
5. Fisher, R. A., 1918. On the correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh* 52: 399–433.
6. Goodman, H. O. and Reed, S. C., 1952. Heredity of fibrosis of the pancreas, possible mutation rate of the gene. *Amer. J. Human. Genet.* 4: 59–71.
7. Haldane, J. B. S., 1936. A search for incomplete sex-linkage in man. *Ann. Eugen.* 7: 28–57.
8. Kempthorne, O., 1957. *An Introduction to Genetic Statistics.* New York, John Wiley and Sons, Inc.
9. Li, C. C., 1961. *Human Genetics.* New York, McGraw-Hill Book Co., Inc.

10. Lowe, C. U., May, C. D., and Reed, S. C., 1949. Fibrosis of the pancreas in infants and children. A statistical study of clinical and hereditary features. *Amer. J. Dis. Child.* 78: 349–374.
11. Merritt, A. D., Hanna, B. L., Todd, C. W., and Myers, T. L., 1963. The advantages of a genetic method in determining the incidence of cystic fibrosis. *J. Am. Med. Assn.* 184: 152.
12. Morton, N. E., 1955. Sequential tests for the detection of linkage. *Am. J. Human Genet.* 7: 277–318.
13. Morton, N. E., 1956. The detection and estimation of linkage between the genes for elliptocytosis and the Rh blood type. *Am. J. Human Genet.* 8: 80–96.
14. Morton, N. E., 1957. Segregation analysis in human genetics. *Science* 127: 79–80.
15. Morton, N. E., 1957 (a). Further scoring types in sequential linkage tests, with a critical review of autosomal and partial sex linkage in man. *Am. J. Human Genet.* 9: 55–75.
16. Morton, N. E., 1959. Genetic tests under incomplete ascertainment. *Am. J. Human Genet.* 11: 1–16.
17. Morton, N. E., 1962. Segregation and linkage. In *Methodology in Human Genetics*, Ed. W. J. Burdette, San Francisco, Holden-Day, Inc.
18. Neel, J. V. and Schull, W. J., 1954. *Human Heredity.* Chicago, Univ. Chicago Press.
19. Renwick, J. H. and Lawler, S. D., 1955. Genetical linkage between the ABO and nail-patella loci. *Ann. Human Genet.* 19: 312–331.
20. Selander, P., 1962. The frequency of cystic fibrosis of the pancreas in Sweden. *Acta Paediat.* 51: 65–67.
21. Smith, C. A. B., 1959. Some comments on the statistical methods used in linkage investigations. *Am. J. Human Genet.* 11: 289–302.
22. Steinberg, A. G., 1959. Methodology in human genetics. *J. Med. Educ.* 34: 315–334.
23. Steinberg, A. G. and Brown, D. C., 1960. On the incidence of cystic fibrosis of the pancreas. *Am. J. Human Genet.* 12: 416–424.
24. Stern, C., 1957. The problem of complete Y-linkage in man. *Am. J. Human Genet.* 9: 147–166.
25. Strandskov, H. H., 1931. A statistical study of the relative goodness of fit of the two proposed theories of human blood group inheritance. *J. Immunol.* 21: 261–277.

Problems of Ascertainment in the Analysis of Family Data*

By James F. Crow, Ph. D., *Department of Medical Genetics, University of Wisconsin, Madison, Wis.*

The study of human genetics has many methodological difficulties. Problems that would be trivially easy in experimental animals and plants require difficult and indirect methods of analysis. Most of the difficulty (except for the obvious obstacle that marriages are not contracted with a view to being maximally informative for genetic analysis) comes from the smallness of families and from the fact that genotypes are often not known. Another problem arises from the fact that the same clinical entity may be inherited in different ways.

There is a special difficulty in distinguishing *familial* from *sporadic* cases. Familial cases are those in which the affected individuals in a sibship have inherited the genes for the trait from one or both parents; sporadic cases are those that are due to new mutations, to chromosome changes, to rare complexes of genes not likely to be repeated in the sibship, or to nongenetic causes. With familial cases, if one child is affected the recurrence risk is 1/4, 1/2 or some other high value depending on the manner of inheritance, penetrance, and viability. With sporadic cases, the recurrence risk is usually very small.

In *Drosophila*, for example, there is no difficulty in distinguishing between a sibship with a 3:1 ratio and one that has a few exceptional flies that arose by mutation or nondisjunction. But in human studies it may be impossible to distinguish between these alternatives. An isolated case in a small sibship may be a familial condition where there happened to be only one affected child; this family, if the number of children had been larger, would have had several affected. On the other hand, the affected child may have been sporadic, in which case there probably would not have been another case even if the family were far beyond the limits of human fertility.

*Paper Number 940 from the Division of Genetics. Supported in part by the National Institutes of Health (RG-8217).

If both parents are heterozygous for a fully penetrant recessive gene, there is the probability $p=1/4$ that any particular child will be affected. If the family has only three children, the probability of no affected child is q^3 and of only one affected is $3pq^2$ where $q=1-p=3/4$. Thus, 27/64 of all three-child families have none affected, and among the 37/64 that have one or more affected, 27/37 have an isolated case. This offers little opportunity to distinguish between familial and sporadic cases, and poor information for testing Mendelian ratios.

In addition to the difficulty with sporadic cases, another troublesome methodological problem arises when a Mendelian ratio is to be compared with the observed ratios in children, where the genotypes of the parents were determined by observation of the children. If sibships of children segregating for a recessive trait are examined, only those families will be included that have one or more affected children, since there is ordinarily no way of knowing in the other cases that the parents were heterozygous. Thus the observed ratio of affected to normal children will be biased.

THE *A PRIORI* CORRECTION

Table 1 gives a set of data on the family distribution of 124 cases of cystic fibrosis with normal parents, arranged according to sibship size. As can be seen, there is a total of 124 affected among 269 children and the affected proportion (46.1 percent) is far in excess of the 25 percent expected on simple recessive inheritance. Clearly, one reason for the excess is the bias previously referred to—the non-inclusion of families where the parents were heterozygous, but failed to have a homozygous recessive child. These would add to the number of normal children and thereby reduce the proportion affected.

This suggests a way of correcting the data. If all families where both parents are heterozygous are included, the expected number of affected children in a sibship of size s is sp. In a family of size s, the probability of having no affected children and thus being omitted from the data is q^s, where $q=3/4$. Thus, in a family of size s, the corrected number of affected children is $sp/(1-q^s)$.

The expected numbers for each family size can be worked out in this way and their total compared with the observed number. This is called the *a priori* method and was introduced by Weinberg (1912). It is described in several textbooks (e.g., Stern, 1960; Neel and Schull, 1954; Li, 1961).

This simple procedure contains a hidden assumption. The method is appropriate only if all families have an equal probability of being included in the sample. In most studies a family with several affected persons has a greater chance of being included than a family with only one affected. As first emphasized in a classical paper by R. A.

Fisher (1934) the manner in which the data are gathered influences the method of analysis. We therefore must look for more generally applicable methods.

For general reviews, see Bailey (1961, Chapter 14) and Morton (1962).

CLASSIFICATION OF ASCERTAINMENT METHODS

I shall use a system of classification and vocabulary introduced by Morton (1959), which follows in some ways an earlier system of Bailey (1951). If the families are obtained by random sampling through the parents without regard to the children (usually not possible for recessive conditions) this is *complete selection*. *Incomplete selection* is ascertainment through the affected children, which means that families with none affected are omitted. With incomplete selection there are three possibilities:

A. *Truncate selection.* Every affected individual is independently ascertained, as if the population were exhaustively examined. A family with a large number of affected children is no more likely to be included than one with a small number. This might happen if ascertainment were by family units. The distribution of affected children among selected families is approximately binomial except for omission of the class with none affected, and thus is truncated.

B. *Single selection.* The method of ascertainment is such that no family is discovered more than once, as would be the case when the probability of ascertainment approaches zero. Such a situation might arise, for example, if the sample were obtained by taking children from a certain grade in school. The probability of a family being included in the study is proportional to the number affected, whereas in truncate selection it is independent of the number affected.

C. *Multiple selection.* Most actual examples are somewhere between these two extremes. If the data are gathered in such a way that the individual ascertainments are identified, or if the ascertainment probability is known, there are appropriate methods of analysis. Multiple selection includes truncate and single selection as special cases.

DEFINITIONS AND SYMBOLS

Any individual in a sibship who is independently ascertained will be called a *proband*. (The word propositus is sometimes used, but since it also has other meanings I shall use proband.)

735-712 O—65——3

Quantities to be estimated:

p=the probability of an affected individual, e.g., $p=1/4$ for a completely penetrant recessive trait with heterozygous parents.

$q=1-p$=the probability of not being affected.

π=probability of an affected person being a proband, *i.e.*; being independently ascertained.

Estimators will be indicated by a circumflex or "hat," e.g. $\hat{p}$.

Observed data:

s=the sibship size.

r=the number of affected persons in the sibship.

a=the number of probands in the sibship.

N=total number of families.

T=total number of children.

R=total number affected.

WEINBERG PROBAND METHOD

The Proband Method was also invented by Weinberg (1912). It has been discussed by Fisher (1934) who provided the error formula. Because of its ease of computation and applicability to all data where the ascertainments have been recorded, it is well suited for an initial analysis. In many cases this will be sufficient, but if a more refined treatment is wanted the Maximum Likelihood Method to be described later can be used.

The argument for the Proband Method is a simple, intuitive one. Each ascertained affected individual (proband) is regarded as providing the information that his parents are capable of producing affected children; then the remaining members of the sibship provide an unbiased estimate of the ratio of affected to normal. If there is more than one proband in the sibship, the family will be counted repeatedly.

A demonstration that the Weinberg Proband Method gives a consistent estimate of the true proportion of recessives is given in appendix 1.

I shall consider multiple selection first, and then regard single selection and truncate selection as special cases.

Multiple Selection

In this procedure we regard a sibship of size s with r affected as a family of $s-1$ with $r-1$ affected. Then each family is counted as many times as there are probands. For example, family 4 in table 1, with 7 children and 3 affected, of which 2 were independently ascer-

tained, would be regarded as a sibship of 6 with 2 affected and would be counted twice. The formula for estimating the true proportion affected, p, is

$$p_0=\frac{\sum a(r-1)}{\sum a(s-1)} \tag{1}$$

where the summation is over all families.

The data in table 1 have been treated in this manner. As can be seen from the table

$$p_0=\frac{59}{218}=.271.$$

TABLE 1.—*Distribution by sibships of 124 individuals with cystic fibrosis (data from Dr. Charles C. Lobeck)*

Sibship number	Number in sibship s	Number affected r	Number of probands a	$a(s-1)$	$a(r-1)$	$a(a-1)$
1	10	3	1	9	2	0
2	9	3	1	8	2	0
3	8	4	1	7	3	0
4	7	3	2	12	4	2
5	7	3	1	6	2	0
6	7	2	1	6	1	0
7	7	1	1	6	0	0
8	6	2	1	5	1	0
9	6	1	1	5	0	0
10	5	3	3	12	6	6
11	5	3	2	8	4	2
12	5	2	1	4	1	0
13	5	2	1	4	1	0
14	5	2	1	4	1	0
15	5	2	1	4	1	0
16	5	2	1	4	1	0
17	5	1	1	4	0	0
18	5	1	1	4	0	0
19	4	3	2	6	4	2
20	4	3	1	3	2	0
21	4	3	1	3	2	0
22	4	2	1	3	1	0
23	4	2	1	3	1	0
24	4	2	1	3	1	0
25	4	2	1	3	1	0
26	4	1	1	3	0	0
27	4	1	1	3	0	0
28	4	1	1	3	0	0
29	4	1	1	3	0	0
30	4	1	1	3	0	0
31	4	1	1	3	0	0
32	3	2	2	4	2	2
33	3	2	2	4	2	2
34	3	2	2	4	2	2
35	3	2	1	2	1	0
36	3	2	1	2	1	0
37	3	2	1	2	1	0
38	3	1	1	2	0	0
39	3	1	1	2	0	0
40	3	1	1	2	0	0
41	3	1	1	2	0	0
42	3	1	1	2	0	0
43	3	1	1	2	0	0
44	3	1	1	2	0	0
45	3	1	1	2	0	0
46	3	1	1	2	0	0
47	3	1	1	2	0	0
48	2	2	2	2	2	2
49	2	2	2	2	2	2
50	2	2	1	1	1	0

TABLE 1.—*Distribution by sibships of 124 individuals with cystic fibrosis (data from Dr. Charles C. Lobeck)*—Continued

Sibship number	Number in sibship s	Number affected r	Number of probands a	$a(s-1)$	$a(r-1)$	$a(a-1)$
51	2	2	1	1	1	0
52	2	2	1	1	1	0
53	2	2	1	1	1	0
54	2	1	1	1	0	0
55	2	1	1	1	0	0
56	2	1	1	1	0	0
57	2	1	1	1	0	0
58	2	1	1	1	0	0
59	2	1	1	1	0	0
60	2	1	1	1	0	0
61	2	1	1	1	0	0
62	2	1	1	1	0	0
63	2	1	1	1	0	0
64	2	1	1	1	0	0
65	2	1	1	1	0	0
66	2	1	1	1	0	0
67	2	1	1	1	0	0
68	2	1	1	1	0	0
69	2	1	1	1	0	0
70	2	1	1	1	0	0
71	2	1	1	1	0	0
72	1	1	1	0	0	0
73	1	1	1	0	0	0
74	1	1	1	0	0	0
75	1	1	1	0	0	0
76	1	1	1	0	0	0
77	1	1	1	0	0	0
78	1	1	1	0	0	0
79	1	1	1	0	0	0
80	1	1	1	0	0	0
	269	124	90	218	59	22

This is the procedure originally suggested by Weinberg. However it assigns too much weight to large families and not enough to small families. In Maximum Likelihood Theory it is shown that optimum weighting is attained by assigning to each group a weight that is inversely proportional to its variance. In practice, this makes only a trivial change in the estimate; but it is only a slight increase in amount of computation and leads at the same time to an estimate of the variance of the estimate.

Fisher (1934) has shown that the variance of the estimate from a single sibship is

$$V=\frac{pq[1+\pi+p\pi(s-3)]}{a(s-1)} \tag{2}$$

where π is estimated by

$$\pi=\frac{\sum a(a-1)}{\sum a(r-1)}. \tag{3}$$

Table 2 shows the data from table 1 rearranged so that the numbers for each sibship are pooled. If each family is to be weighted in inverse proportion to its variance, this can be accomplished by multiplying each family size estimate by C, where $1/C=1+\pi+p_0\pi(s-3)$.

TABLE 2.—*Data from table 1 grouped by sibship size*

s	$a(s-1)$	$a(r-1)$	$a(a-1)$	$1/C$	$Ca(s-1)$	$Ca(r-1)$
10	9	2	0	2.08	4.33	.963
9	8	2	0	1.98	4.04	1.010
8	7	3	0	1.88	3.73	1.597
7	30	7	2	1.78	16.88	3.940
6	10	1	0	1.68	5.97	.597
5	48	15	8	1.58	30.40	9.530
4	42	12	2	1.47	28.45	8.130
3	38	9	6	1.37	27.65	6.550
2	26	8	4	1.27	20.40	6.290
	218	59	22	----------	141.85	38.607

$p_0=\frac{59}{218}=.271$ (first estimate)

$\hat{\pi}=\frac{22}{59}=.373$

$\frac{1}{C}=1+\hat{\pi}+p_0\hat{\pi}(s-3)$

$\hat{p}=\frac{38.607}{141.85}=.272$ (improved estimate)

$V_{\hat{p}}=\frac{\hat{p}(1-\hat{p})}{141.85}=.001396$

$\sigma_{\hat{p}}=\sqrt{.001396}=.037$

The improved estimate, $\hat{p}$, is given by

$$\hat{p}=\frac{\sum Ca(r-1)}{\sum Ca(s-1)} \tag{4}$$

where a, r, and s now stand for the pooled numbers in all sibships of the same size. The variance of $\hat{p}$ is estimated by

$$V_{\hat{p}}=\frac{\hat{p}(1-\hat{p})}{\sum Ca(s-1)} \tag{5}$$

where

$$C=\frac{1}{1+\pi+p_0\pi(s-3)}$$

The data of table 2 are analyzed in this manner. As can be seen, the improved estimate of p is hardly different from the first estimate. From these data, the best estimate of the true proportion of affected is .272 with a standard error of .037. The absence of any difference in the first and the improved estimates shows that there is no consistent association between proportion affected and sibship size, although this could of course be tested more directly.

Single Selection ($\pi \to 0$)

As the probability of ascertainment gets small it becomes less and less likely that a family will be ascertained more than once. In the

limit $a=1$ and $C=1$ for any ascertained family, and the proportion affected is estimated by

$$\hat{p}=\frac{\sum(r-1)}{\sum(s-1)}=\frac{R-N}{T-N}. \tag{6}$$

The variance becomes

$$V_{\hat{p}}=\frac{\hat{p}\hat{q}}{\sum(s-1)}=\frac{\hat{p}\hat{q}}{T-N}. \tag{7}$$

This estimate is a maximum likelihood estimate (Fisher, 1934; Neel and Schull, 1954), hence is fully efficient. Thus the proband method increases in efficiency as π gets smaller.

Notice that if this is applied to the data of table 1, $\hat{p}=\frac{124-80}{269-80}=.233$. That this underestimates the true proportion is expected since it is applied to data where some families were ascertained more than once.

The variance of this estimate, which would be appropriate if the ascertainment probability were very small, is $(.233)(.767)/189=.000945$. The standard error is .031, so in this example the two estimates, one based on an ascertainment probability of 0 and other with $\pi=.37$, do not differ significantly from each other.

Truncate Selection ($\pi=1$)

In this case every affected individual is ascertained independently, so $a=r$. The estimate of p is $p_0=\frac{\sum r(r-1)}{\sum r(s-1)}$, or if the data are grouped by family size as in table 2, $\hat{p}=\frac{\sum Cr(r-1)}{\sum Cr(s-1)}$. The variance formula is the same as that given for multiple selection—Equations (2) and (5)—when $\pi=1$ and $a=r$.

Applied to the data of table 1, we obtain $p_0=112/363=.309$. This is a substantial overestimate because the formula for truncate selection was applied to data where the ascertainment probability was in fact considerably less than 1.

The three estimates, .233 when π is assumed to be zero, .272 when π is .37, and .309 when π is assumed to be one, illustrate the dependence of p on the ascertainment probability. Clearly, the estimates differ; but perhaps the most important point is that the estimate of p, though somewhat dependent on π, is not very sensitive to errors in the estimation of π.

In data where the ascertainments are not recorded and when there is no other basis for estimating π, about the only procedure available is to analyze the data on the two extreme assumptions of $\pi=0$ and

$\pi=1$, and take the estimate of p to be at some unknown point between the two extreme estimates.

The ascertainment probability, π, though needed for the estimation of p, is of no genetic interest. However, the number of ascertainments divided by π gives an estimate of the prevalence of the condition (see Morton, 1962).

THE HALDANE METHOD

If $\pi=1$, that is with truncate selection, Fisher's Maximum Likelihood Method is available, although the computations are more difficult. The procedure was first used by Haldane (1932). The proportion affected, p, is estimated from the total number affected, R, by $\hat{p}$, in the equation

$$R=\sum_{s=2}^{\infty}\frac{sn_s\hat{p}}{1-\hat{q}^s} \tag{8}$$

where n_s is the number of sibships of size s, and $\hat{q}=1-\hat{p}$. In this case, R is the total number affected in families of size two or more; families of one are not informative and are therefore omitted. The variance of $\hat{p}$ is given by

$$\frac{1}{V_{\hat{p}}}=\sum_{s=2}^{\infty}\frac{sn_s}{\hat{p}\hat{q}}\left[\frac{1-s\hat{p}\hat{q}^{s-1}-\hat{q}^s}{(1-\hat{q}^s)^2}\right]. \tag{9}$$

Equation (8) is most easily solved by trial and error, choosing trial values of p and estimating $\hat{p}$ by interpolation. The calculations are greatly simplified by the use of tables. Li (1961, p. 66) gives tables of $sp/(1-q^s)$ and $s(1-spq^{s-1}-q^s)/pq(1-q^s)^2$ for various values of p and s.

Using Li's tables and the data in table 1, the expected total number is 101.0 when $p=.25$, 108.6 when $p=.30$ and 115.8 when $p=.35$. The observed value, R, is 115. Interpolating, $\hat{p}=.355$. That this overestimates the value of p is expected since the assumption of truncate selection is inappropriate for these data.

If π is other than 1, but known, the Maximum Likelihood Method can be extended to include this (Haldane, 1938). The computation is troublesome, but the procedure is discussed by Neel and Schull (1954, p. 225). I shall not dissuss it because it is very much like the methods of the next section.

IMPROVED PROCEDURES

For routine purposes the Weinberg Proband Method is easy to use and generally satisfactory. It is, I believe, the method of choice for a first analysis of the data. There is the further advantage that the

more elaborate Maximum Likelihood Method usually requires an iterative solution based on a trial value of p; for this purpose the value estimated by the Proband Method can be used.

There are two principal criticisms of the methods discussed thus far. One is that, although these provide estimates of p based on certain assumptions, there are no tests for the validity of the assumptions. The second objection is that these provide no basis for separation of familial and sporadic cases.

For more extensive tests Fisher's Maximum Likelihood Method is the standard procedure. This permits more elaborate hypotheses to be tested and allows for internal checks. The difficulty is that the equations can almost never be solved directly, so that complex iterative methods are needed. The calculations are greatly facilitated by the use of scoring procedures developed by Fisher. These have been adapted for segregation analysis by Finney (1949), Bailey (1951, 1961), and Morton (1959).

The adaptation of maximum likelihood scores for separation of familial and sporadic cases and for tests of the assumptions are mainly due to Morton (1959). I shall illustrate some of the procedures by a further analysis of the cystic fibrosis data already discussed.

The first problem is the separation of sporadic cases (mutations, phenocopies, etc.) from familial cases. This can be accomplished by distinguishing between multiplex sibships (two or more affected) and simplex (only one affected). If both sporadic and familial cases are rare and uncorrelated in occurrence, and if the recurrence rate is very small for sporadic cases, then multiplex sibships will consist almost exclusively of familial cases. So the operational procedure is this: We assume that any sibship with two or more affected is familial, whereas simplex sibships are a mixture of sporadic cases and familial sibships which happen to have only one affected. We therefore analyze multiplex and simplex families separately and attribute any excess of affected in the latter to sporadic cases.

Letting x be the proportion of sporadic cases among all cases, we can estimate x and p from the distribution of simplex *vs.* multiplex families. We can get an independent estimate of p from the distribution of the number affected in multiplex families.

The parameters, such as p, x, and π, are estimated by the Method of Maximum Likelihood. In this method the probability of the observed data is expressed as a function of these parameters, and the values of the parameters that jointly maximize the probability of the observed results are taken as the best estimates. This is accomplished by equating the derivative of the probability to zero and solving the resulting equations. One improvement is to use the logarithm of the probability of the observed results rather than the probability; this simplifies the algebra and doesn't change the result, for the same values of the parameters that maximize the probability also maximize

the log of the probability. The log-probability expression also lends itself to a procedure for finding the solutions by iteration.

We start with a trial value of the quantity to be estimated, say p. The Weinberg Sib Method provides a good starting value. Calling the trial value p_0, we obtain an improved estimate, p_1, by the equation

$$p_1=p_0+\frac{U_{p_0}}{K_{p_0p_0}} \tag{10}$$

where U_{p_0} and $K_{p_0p_0}$ are quantities that will be defined later. From this improved estimate a still better estimate, p_2, is given by

$$p_2=p_1+\frac{U_{p_1}}{K_{p_1p_1}} \tag{11}$$

and so on, until the desired accuracy is attained.

Each family is assigned a score, u_p, and contributes an amount of information, k_{pp}, relative to the quantity p to be estimated. These quantities are additive, so the total score for all families is $U_p=\Sigma u_p$ and $K_{pp}=\Sigma k_{pp}$.

The score, u_{p_0}, is the derivative of the logarithm of the probability with respect to p, evaluated at the trial value p_0. Likewise, the amount of information is the expected value of the second derivative, also evaluated at the trial value.

The general theory of maximum likelihood scores is discussed in statistics textbooks (e.g., Rao, 1952). For a general discussion of maximum likelihood scoring methods as applied to problems in segregation analysis see Bailey (1961, appendix 1).

I shall illustrate some of Morton's simpler methods by continuing the analysis of the data in table 1. To avoid any contribution from sporadic cases, I shall consider first only multiplex sibships ($r>1$). Within such sibships the probability of r affected in a sibship of size s, and of the sibship being ascertained at least once is

$$P=\frac{\binom{s}{r}p^rq^{s-r}[1-(1-\pi)^r]}{1-(1-p\pi)^s-\pi spq^{s-1}}. \tag{12}$$

The estimate of π can be obtained conveniently in either of two ways. The distribution of probands is clearly equivalent to truncate selection, since (1) there are no families represented unless there was at least one proband, and (2) there is no ascertainment bias in proband selection—by the mere fact of their being probands they are independently ascertained.

Thus π can be estimated (as it was earlier in the Weinberg method) by $\frac{\sum a(a-1)}{\sum a(r-1)}$ or by the Haldane Method (replacing R by the total number of probands, s by the number affected, r, and n_s by the number of

families with r affected in equation (8)). These two procedures give almost identical results, .37 and .36, as estimates of π.

The score to be assigned to each sibship is

$$u_p=\frac{r}{pq}-e_p \tag{13}$$

where values of e_p are given for various family sizes in table 3 for the trial value of $p_0=.25$. Likewise, table 3 gives the information values, k_{pp}, to be assigned to a sibship of size s. The values given in table 3 are based on $\pi=.35$.

Each multiplex family in table 1 is given a u and k score. For example, in sibship number 1 where $s=10$ and $r=3$

$$u_p=3(5.3333)-17.192=1.192$$
$$k_{pp}=34.873.$$

Adding up the values for the 27 informative sibships (those with at least three sibs and with two or more affected) gives

$$U_p=\sum u_p=-0.350$$
$$K_{pp}=\sum k_{pp}=298.393.$$

The improved estimate of p is

$$p_1=.25-.350/298.39=.249.$$

A better estimate could be obtained by repeating the procedure with a trial value of .249. This would necessitate a new table 3 based on this value. However, the first estimate was so close that no further iteration is necessary. The variance of the estimate is the reciprocal of K_{pp}, hence

$$V_{p_1}=1/298.39=.00335$$

and the standard error of p is

$$\sigma_{p_1}=\sqrt{V_{p_1}}=.058$$

Now, as an independent estimate of p and as a test for sporadic cases, we compare the simplex and multiplex families. Among the 71 informative sibships (those with at least 2), there are 38 simplex and 33 multiplex. Each family is assigned a score depending on only two things, the family size and whether it is simplex or multiplex. Such scores are given in table 4, based on the values $p=.25$, $\pi=.35$, and $x=0$.

For example, family 1 is multiplex ($r>1$) with 10 sibs; therefore its scores are: $u_p=1.299$, $u_x=-1.459$, $k_{pp}=13.72$, $k_{xx}=17.30$, and $k_{px}=-15.41$. Family 7, being simplex with 7 sibs, is scored as: $u_p=-6.989$, $u_x=4.324$, $k_{pp}=14.62$, $k_{xx}=5.60$, and $k_{px}=-9.05$.

Proceeding in this manner through all 71 families and adding all the scores together, we obtain for the total scores:

$$U_p = 16.026 \qquad K_{pp} = 642.90$$
$$U_x = -7.848 \qquad K_{xx} = 125.95$$
$$K_{px} = -260.49.$$

When two or more parameters, such as p and x in this case, are being estimated simultaneously, it is conventional to write the K's in the form of an Information Matrix:

$$\begin{vmatrix} K_{pp} & K_{px} \\ K_{px} & K_{xx} \end{vmatrix} = \begin{vmatrix} 642.90 & -260.49 \\ -260.49 & 125.95 \end{vmatrix}$$

To obtain the variances and the improved estimates, the matrix is inverted. This is easily done for small matrices, such as the 2 x 2 matrix of this example. We designate by K^{pp} the element in the inverted matrix corresponding to the position of the element K_{pp} in the original matrix, and so on. Then the elements of the inverted matrix are $K^{pp} = K_{xx}/\Delta$, $K^{xx} = K_{pp}/\Delta$, and $K^{px} = -K_{px}/\Delta$, where $\Delta = K_{pp}K_{xx} - (K_{px})^2$. Thus, the inverted matrix (called the Variance-Covariance matrix) is

$$\begin{vmatrix} K^{pp} & K^{px} \\ K^{px} & K^{xx} \end{vmatrix} = \begin{vmatrix} .00960 & .01986 \\ .01986 & .04901 \end{vmatrix}$$

The improved estimates for the two parameters are

$$p_1 = p_0 + U_{p_0}K^{p_0p_0} + U_{x_0}K^{p_0x_0}$$
$$x_1 = x_0 + U_{x_0}K^{x_0x_0} + U_{p_0}K^{p_0x_0} \qquad (14)$$

and the variances are

$$V_{p_1} = K^{p_0p_0}, \; V_{x_1} = K^{x_0x_0}.$$

In this example, $p_0 = .25$ and $x_0 = 0$. Hence

$$p_1 = .25 + (16.026)(.00960) + (-7.848)(.01986) = .248$$
$$V_{p_1} = .00960, \; \sigma_{p_1} = .098$$
$$x_1 = 0.00 + (-7.848)(.04901) + (16.026)(.01986) = -.066$$
$$V_{x_1} = .04901, \; \sigma_{x_1} = .221.$$

Again, the trial value of p was close enough so that a second cycle of iteration is not needed. In these data there is excellent agreement with the .25 proportion of recessive homozygotes expected on Mendelian theory. Furthermore, there is no evidence for sporadic cases,

since x does not deviate significantly from zero. On the other hand, the standard error of the estimate of x is so large that even a substantial number of sporadic cases cannot be ruled out.

One of the most convenient properties of the u and k scores is that they are additive for different sets of data. We can combine the results of the two independent analyses by summing the scores. Thus, the combined scores are:

$$U_p = -.350 + 16.026 = 15.676$$

$$K_{pp} = 298.393 + 642.90 = 941.293.$$

These lead to the new information matrix

$$\begin{vmatrix} 941.29 & -260.49 \\ -260.49 & 125.95 \end{vmatrix}$$

which inverts to the variance-covariance matrix

$$\begin{vmatrix} .00248 & .00514 \\ .00514 & .01857 \end{vmatrix}.$$

The combined estimates from both analyses are

$$\hat{p} = .25 + (15.676)(.00248) + (-7.848)(.00514) = .2485$$

$$V_{\hat{p}} = .00248, \quad \sigma_{\hat{p}} = .050$$

$$\hat{x} = 0.00 + (-7.848)(.01857) + (15.676)(.00514) = -.065$$

$$V_{\hat{x}} = .0186, \quad \sigma_{\hat{x}} = .136.$$

Clearly, in these data there is excellent agreement with the 3:1 Mendelian expectation and no evidence for any appreciable contribution from sporadic cases. In fact, taken at face value, the number of isolated cases is slightly less than expected, as indicated by the negative sign of U_x. However, the occurrence of a substantial fraction of sporadic cases cannot be ruled out because of the large variance of the estimate of x.

Although sporadic cases do not appear to be an important part of the total incidence in cystic fibrosis, this is not true of many other conditions. In deafness there is a large component caused by recessive genes, but also a substantial sporadic fraction, as revealed by this kind of analysis. For the details see Chung, Robison and Morton (1959).

It may appear puzzling that the Maximum Likelihood estimate of p has a larger variance (.00248) than the variance from the Proband Method (.001396). This is because I assumed in the Proband analysis that there were no isolated cases, whereas in the Maximum Likelihood this was not assumed. If we make this assumption, then the variance of p is simply the reciprocal of K_{pp}, $1/941.29 = .00106$. The estimate of p is $\hat{p} = .25 + (15.676)(.00106) = .267$.

We can compare these results with those obtained by the Proband Method:

Proband: $\hat{p}=.272$, $V_{\hat{p}}=.00140$, $\sigma_{\hat{p}}=.037$

Maximum Likelihood: $\hat{p}=.267$, $V_{\hat{p}}=.00106$, $\sigma_{\hat{p}}=.033$.

The greater efficiency of the Maximum Likelihood Method is shown by its smaller variance. The relative efficiency of the Proband Method may be expressed as the ratio of the variances, 106/140=76 percent. In other words 76 observations analyzed by Maximum Likelihood provides as much statistical precision as 100 observations analyzed by the Proband Method.

In all the foregoing discussion I have not taken into consideration errors in estimating the ascertainment probability, π. We estimated π from the data, but since π enters directly into the computation of p, errors in the estimation of π will affect the estimate of p. This can be taken into account in the Maximum Likelihood Method by including u_π, $k_{\pi\pi}$, $k_{p\pi}$, and $k_{z\pi}$ with the other scores (Morton and Chung, 1959). The effect will be to increase $V_{\hat{p}}$, so the procedures that I have used have both overestimated the precision of the estimate, $\hat{p}$. However, the effect of this is not great, for as pointed out earlier the estimate of p is relatively insensitive to small errors in the estimation of π.

This kind of analysis can be carried considerably further. Additional tests for internal consistency can be included. The distribution of the number of times that an individual proband is ascertained provides a test for the assumption of the independence of separate ascertainments (Morton, 1959). The same kind of analysis can also be extended to other kinds of sibships to test dominant and sex-linked inheritance. For other examples and more details see Morton (1959), Morton and Chung (1959), and Morton (1962).

Tables of u and k scores, of which tables 3 and 4 are examples, are available for various values of the parameters. Also the procedures have been adapted to computer analysis; the program is called SEGRAN (Morton, 1962).

TABLE 3.—*Scores for distribution of number affected* (r) *in sibships with more than one affected* (r>1)

$p=\frac{1}{4}$ $\pi=.35$ $\frac{1}{pq}=5.3333$ $u_p=\frac{r}{pq}-e_p$

s	e_p	k_{pp}
3	11.320	3.057
4	12.022	6.540
5	12.773	10.431
6	13.571	14.704
7	14.414	19.324
8	15.300	24.254
9	16.227	29.451
10	17.192	34.873

TABLE 4.—*Scores for simplex* (r=1) vs. *multiplex* (r>1) *sibships*

$p=\frac{1}{4}$ $x=0$ $x=.35$

s	$r=1$		$r>1$		k_{pp}	k_{xx}	k_{px}
	u_p	u_x	u_p	u_x			
2	−1.150	.288	4.183	−1.046	4.81	.30	−1.20
3	−2.306	.685	3.680	−1.093	8.49	.75	−2.52
4	−3.468	1.229	3.220	−1.141	11.17	1.40	−3.96
5	−4.636	1.970	2.803	−1.191	13.00	2.35	−5.52
6	−5.810	2.972	2.428	−1.242	14.11	3.69	−7.22
7	−6.989	4.324	2.092	−1.294	14.62	5.60	−9.05
8	−8.174	6.144	1.793	−1.348	14.66	8.28	−11.02
9	−9.364	8.586	1.530	−1.403	14.33	12.04	−13.14
10	−10.560	11.859	1.299	−1.459	13.72	17.30	−15.41

ADDITIONAL INFORMATION ON THE FREQUENCY OF ISOLATED CASES

If the trait being analyzed is caused by one or more rare recessive genes there will be an appreciable contribution to familial cases from consanguineous marriages. However the sporadic cases are probably not increased appreciably by inbreeding. This provides another basis for separation of sporadic from familial cases.

Following Morton (1962 and earlier papers) we let

y =proportion of sporadic among isolated cases

α =average inbreeding coefficient in the population

F_i=inbreeding coefficient of isolated cases

F_f=inbreeding coefficient of familial cases;

then, since simplex families are a mixture of sporadic and chance isolated cases,

$$F_i=y\alpha+(1-y)F_f$$

from which

$$y=\frac{F_f-F_i}{F_f-\alpha} \tag{15}$$

and

$$x=cy \tag{16}$$

where c is the proportion of simplex families. In obtaining c and F_f the families are weighted by the number of probands.

The data on cystic fibrosis contain so few instances of parental consanguinity that this method is of no use. It was used very effectively by Chung, Robison, and Morton (1959) in the analysis of deaf mutism. The proportion of sporadic cases, x, as estimated by equation (14) agreed very well with the completely independent estimate obtained from equations (15) and (16).

ESTIMATION OF THE NUMBER OF LOCI INVOLVED

It is often difficult to distinguish between a situation where all cases of a disease are caused by the same gene and the alternative where any one of several genes can cause the condition. Sometimes this distinction can be made clinically; in other instances there may be different modes of inheritance. In rare cases linkage can be used. For example, Morton (1956) showed that linkage of elliptocytosis with the Rh locus is significantly heterogeneous; the data are most satisfactorily explained by two clinically indistinguishable loci, one closely linked to the Rh locus and the other independent of it. In practice this method is almost completely restricted to dominant or X-linked conditions.

For recessive traits it is desirable to know whether all cases are caused by the same gene or whether any one of several genes, when homozygous, can produce the condition. One approach is through consanguinity data.

If the trait is completely recessive its frequency in a population with inbreeding coefficient F is $(1-F)\Sigma q_i^2+F\Sigma q_i$, where q_i is the frequency of the individual recessive allele. The larger the number of genes, the smaller will be the frequency of each, and Σq_i will be correspondingly larger than Σq_i^2. Since, with fully penetrant recessives Σq_i^2 is the trait frequency in a randomly mating population, the increase with inbreeding relative to the normal incidence can be used to measure the number of genes.

The regression of the incidence (after sporadic cases are removed) on the inbreeding coefficient F may be written as

$$I=1-e^{-(A+BF)}\approx A+BF \qquad (17)$$

where $A+B=\Sigma q_i=n\bar{q}$ and $A=\Sigma q_i^2>(n\bar{q})^2$ from which $(A+B)^2/A$ gives a minimum estimate of the number of genes (Morton, 1962 and earlier). This is a correct estimate only in the unlikely circumstance that all alleles are equally frequent, otherwise it is too small. This procedure, applied by Chung, Robison, and Morton (1959) to data on deaf mutism led to the estimate of at least 36 alleles.

It should be mentioned that this method does not tell whether these alleles are at different loci. Formally, they could all be at the same locus provided all heterozygous combinations had normal hearing. The hypothesis of many separate loci is more consistent with evidence from experimental organisms.

The second procedure follows a suggestion of A. G. Steinberg (1962). We inquire into the frequency of the trait in relatives (usually cousins) of index cases. If the gene is rare the frequency of the trait in cousins will be nearly proportional to the gene frequency in the population, the proportionality constant being the

probability that the cousin carries the same allele as the homozygous propositus. If the trait can be caused by any of several recessive genes their individual frequency is less and therefore the incidence in cousins of probands will be less.

We can derive a quantitative expression for cousins by utilizing figure 1.

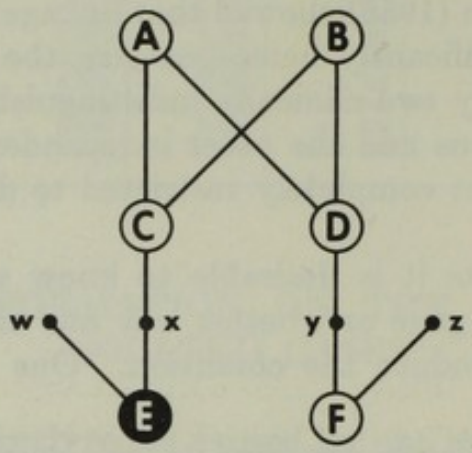


FIGURE 1.—A pedigree showing two cousins. Large letters designate persons; small letters, gametes. Gametes z and w come from the unrelated spouses not shown in the pedigree.

We desire the probability that F is affected, given that E is affected. If E is affected either A or B must be a carrier, unless there has been a mutation, which possibility we ignore. The probability that y carries the same allele as x, having inherited it from the same grandparent, is 1/4. There is also a chance that both grandparents, A and B, were carriers, but if the recessive gene frequency is small this probability is negligible. The chance that gamete z carries the recessive gene is simply p, the population gene frequency. Thus the probability of an affected cousin is

$$P=\frac{p}{4} \text{ approximately;} \tag{18}$$

the population incidence, I, is

$$I=p^2. \tag{19}$$

If there are k genes, of equal frequency, p then approximately

$$P=\frac{\bar{p}}{4} \tag{20}$$

$$I=\Sigma p^2=k\bar{p}^2 \tag{21}$$

and k can be estimated by eliminating $\bar{p}$ from (20) and (21)

$$k=\frac{I}{16P^2} \text{ approximately.} \tag{22}$$

A more accurate formula is derived in Appendix 2, but is not needed for any data available at present. It is important to emphasize that any variation in the frequency of p from locus to locus will decrease the value of k, so equation (22) will tend to underestimate the number of loci if k is greater than 1.

Dr. A. G. Steinberg (personal communication) reports some data from Ohio giving $I=.000267$ and $P=.00425$. Taking these figures at face value gives $k=(.000267)/16(.00425)^2=.92$. Thus there is no evidence for more than one locus. However, there is considerable uncertainty, especially in the incidence figure. Until better data are available, this should be regarded only as illustrative of the method.

ACKNOWLEDGMENTS

I should like to thank Dr. Charles C. Lobeck for supplying the unpublished cystic fibrosis data used throughout the article. I am indebted to Dr. George Hutchinson for pointing out a better weighting system in the Proband Method than I had previously used. Dr. Arthur Steinberg has been most helpful in discussions of the problem of estimating the number of genes. Mr. Takeo Maruyama prepared the computer program for the Maximum Likelihood scores in tables 3 and 4, which are based on some earlier tables of Dr. Newton Morton.

SUMMARY

When the segregation ratio for a recessive gene is estimated from sibships where the heterozygosity of the parents has been ascertained through their having produced one or more affected children, the observed proportion of homozygous recessive individuals generally overestimates the true proportion. Several methods for eliminating this bias are described. The Weinberg Proband Method is simple to use and leads to a consistent estimate of the segregation ratio, but it is not fully efficient (in the statistical sense). It is useful when the greatest of precision is not needed, and as a first estimate for iterative approaches to more precise estimates. The Maximum Likelihood Method is asymptotically efficient and has the advantage of permitting tests of some of the assumptions and separation of familial from sporadic cases (mutations, phenocopies, etc.) in sibships where there is only one affected person. The methods are described and illustrated with data from sibships where at least one child has cystic fibrosis of the pancreas.

The fraction of sporadic cases can also be estimated when the gene is rare by the difference in parental consanguinity rate between sibships with only one affected (simplex) and those with more than one (multiplex).

735-712 O—65——4

Two ways are described for determining whether all cases of a disease are caused by the same recessive gene, or any of several genes may cause the disease. One method depends on the relative incidence in children of related and unrelated parents. The other depends on the incidence in the cousins of index cases.

APPENDIX 1

DEMONSTRATION THAT THE PROBAND METHOD GIVES A CONSISTENT ESTIMATE OF THE SEGREGATION RATIO

Consider sibships of size s. Let x be the probability that a family with r affected will be ascertained at least once. Then

$$x=1-(1-\pi)^r.$$

Let a be the number of ascertainments in a family of size r. The estimated number of ascertainments among families ascertained at least once is

$$a=\frac{\pi r}{1-(1-\pi)^r}=\frac{\pi r}{x}.$$

The value yielded by the sib method, $\hat{p}$, is

$$\hat{p}=\frac{\sum_{r=1}^{s} a(r-1)x\binom{s}{r}p^r q^{s-r}}{\sum_{r=1}^{s} a(s-1)x\binom{s}{r}p^r q^{s-r}}$$

where p is the true probability of an affected individual. Replacing a by its estimated value, $\hat{a}=\pi r/x$,

$$\hat{p}=\frac{\pi\sum_{r=1}^{s} r(r-1)\binom{s}{r}p^r q^{s-r}}{\pi(s-1)\sum r\binom{s}{r}p^r q^{s-r}}$$

$$=\frac{s(s-1)p^2\sum_{r=2}^{s}\frac{(s-2)!}{(r-2)!(s-r)!}p^{r-2}q^{s-r}}{s(s-1)p\sum_{r=1}^{s}\frac{(s-1)!}{(r-1)!(s-r)!}p^{r-1}q^{s-r}}$$

$$=\frac{s(s-1)p^2(p+q)^{s-2}}{s(s-1)p(p+q)^{s-1}}$$

$$=p \text{ since } p+q=1.$$

When $\pi\to 0$, $x\to\pi r$, and $a\to 1$, then

$$\hat{p}=\frac{\sum\pi r(r-1)\binom{s}{r}p^r q^{s-r}}{\sum\pi r(s-1)\binom{s}{r}p^r q^{s-r}}$$

which is the same expression as before.

APPENDIX 2

A MORE EXACT FORMULA FOR THE INCIDENCE OF A RECESSIVE DISEASE IN THE COUSINS OF PROBANDS

Referring to Figure 1, if E is affected then at least one of the grandparents, A and B, must be a carrier (unless there has been a mutation, which possibility we ignore). The *a priori* probability of both A and B being carriers is $(2pq)^2$, where p is the frequency of the abnormal gene and $q=1-p$ is the frequency of the normal gene. The probability that one or the other is a carrier, but not both, is $2\times 2pq \times q^2=4pq^3$. Thus the ratio of the two probabilities is $4p^2q^2/4pq^3$, or $p:q$. We then use Bayes' principle as follows:

	A or B is a carrier	Both A and B carriers
A priori probability ratio	q	p.
Conditional probability that C and D are both normal.	1	$(3/4)^2$.
Conditional probability that x carries an abnormal gene.	1/4	1/3 (1/3 rather than 1/2 because C is normal).
A posteriori probability ratio	$q/4=X$	$3p/16=Y$.

Therefore the probability that gamete y carries the recessive gene is

$$\frac{1}{4}\frac{X}{X+Y}+\frac{1}{3}\frac{Y}{X+Y}=\frac{1}{4(1-p/4)}\approx\frac{1}{4}+\frac{p}{16}$$

The probability that gamete z carries the abnormal gene is p, the population gene frequency. (This is p rather than $p/(1+p)$ because the relevant gene frequency is the frequency at birth, not in adulthood.)

Putting all this together, the probability that individual F is affected is

$$P=\frac{p}{4}+\frac{p^2}{16}, \text{ or approximately } P=p/4$$

(This ignores the information provided by possible additional normal sibs of C and D, which would reduce the $p^2/16$ component slightly.)

If there are k genes at different loci, any one of which can cause the disease when homozygous, the probability of an affected cousin is

$$P=\frac{1}{4}\Sigma x_i p_i\left(1+\frac{p_i}{4}\right)+\Sigma(1-x_i)p_i^2 \qquad (1)$$

$$x_i=\frac{p_i^2}{\Sigma p_i^2} \qquad (2)$$

$$I=\Sigma p_i^2 \qquad (3)$$

where I is the incidence of the disease and x_i is the proportion of cases caused by the i^{th} gene. The right term in equation (1) comes from cases in F due to a different gene than the one causing the disease in E.

If the k genes are equally frequent, neglecting $p^2/16$, we have

$$P=\frac{1}{4}p+(k-1)p^2 \qquad (4)$$

$$I=kp^2 \qquad (5)$$

from which k can be estimated from P and I by eliminating p.

This leads to

$$k=\frac{I}{(P-cI)^2} \tag{6}$$

where $c-(k-1)/k$. This can be most easily solved by using equation (22) to give a trial value for k. This value is used to calculate c, which is substituted into equation (6) to give an improved estimate of k. The process may be repeated if necessary for greater accuracy. The improved estimate for Steinberg's data is 0.93, hardly different from the value 0.92 obtained by equation (22). If the genes are of unequal frequency this will generally be an underestimate of k.

REFERENCES

1. Bailey, N. T. J., 1951. A classification of methods of ascertainment and analysis in estimating the frequencies of recessives in man. *Ann. Eugen.* 16: 223–225.
2. Bailey, N. T. J., 1961. *Mathematical Theory of Genetic Linkage.* Oxford: The Clarendon Press.
3. Chung, C. S., Robison, O. W., and Morton, N. E., 1959. A note on deaf mutism. *Ann. Human Genet.* 23: 357–366.
4. Finney, D. J., 1949. The truncated binomial distribution. *Ann. Eugen.* 14: 319–328.
5. Fisher, R. A., 1934. The effects of methods of ascertainment upon the estimation of frequencies. *Ann. Eugen.* 6: 13–25.
6. Haldane, J. B. S., 1932. A method for investigating recessive characters in man. *J. Genet.* 28: 251–255.
7. Haldane, J. B. S., 1938. The estimation of the frequencies of recessive conditions in man. *Ann. Eugen.* 8: 255–262.
8. Li, C. C., 1961. *Human Genetics: Principles and Methods.* New York: McGraw-Hill Book Co.
9. Morton, N. E., 1956. The detection and estimation of linkage between the genes for elliptocytosis and the RH blood type. *Amer. J. Hum. Genet.* 8: 80–96.
10. Morton, N. E., 1959. Genetic tests under incomplete ascertainment. *Amer. J. Hum. Genet.* 11: 1–16.
11. Morton, N. E., 1962. Segregation and linkage. In *Methodology in Human Genetics,* Ed. by W. Burdette. San Francisco: Holden-Day, Inc., pp. 17–52.
12. Morton, N. E., and Chung, C. S., 1959. Formal genetics of muscular dystrophy. *Amer. J. Human Genet.* 11: 360–379.
13. Neel, J. V., and Shull, W. J., 1954. *Human Heredity.* University of Chicago Press.
14. Rao, C. R., 1952. *Advanced Statistical Methods in Biometric Research.* New York: John Wiley and Sons.
15. Steinberg, A. G., 1962. Population genetics: Special cases. *Methodology in Human Genetics,* Ed. by W. Burdette. San Francisco: Holden-Day, Inc., pp. 76–91.
16. Steinberg, A. G. Personal communication.
17. Stern, C., 1960. *Principles of Human Genetics.* San Francisco: W. H. Freeman and Co.
18. Weinberg, W., 1912. Methode und Fehlerquellen der Untersuchung auf Mendleschen Zahlen beim Menschen. *Arch. Rass. u. Ges. Biol.* 9: 165–174.

Estimation of Genetic Parameters in Population Studies

By William J. Schull, Ph. D., *Department of Human Genetics, School of Medicine, University of Michigan, Ann Arbor, Mich.*

Interwoven into the fabric of many, if not all, epidemiologic studies are a variety of parameters of genetic interest. Among these are or may be:

1) the frequencies of the genes associated with the characteristics under study assuming, of course, that these characteristics are either genetic in origin or are significantly altered in their manifestation by one's genetic constitution;

2) the relative fitnesses to be assigned to each of the genotypes; that is, the extent to which individuals of a given genotype, for all genotypes, are represented by progeny in subsequent generations as contrasted with some standard; and

3) the rates of spontaneous mutation.

On occasion, one may be interested in

4) the frequency of nonpaternity, *i.e.*, the frequency with which the mother of a child incorrectly identifies the child's father;

5) the correlation between uniting gametes either as a consequence of inbreeding or of phenotypic assortative mating; and

6) the existence of maternal and/or paternal effects.

Needless to say, other important parameters exist and it should be patent that, brief as this enumeration is, human genetics does not want for values to estimate. It should be equally obvious that a presentation such as this can hardly be exhaustive. Accordingly, we shall direct our attention toward only two problems, and even then the treatment cannot be complete. Specifically, we shall consider the estimation of gene frequencies in panmictic, that is, in randomly breeding populations and the estimation of the average coefficient of inbreeding. We have chosen the first of these because of the commonness of its occurrence; the second has been chosen largely because

of the current interest in inbreeding as a means of evaluating genetic loads—the burdens imposed upon a population by virtue of the existence of individuals not of the optimal type (see Crow, 1963; Levene, 1963; Li, 1963; Sanghvi, 1963). Before we turn to these problems, however, certain general remarks on the estimation of genetic parameters seem in order.

Some general remarks on the estimation of genetic parameters: Rarely, if ever, are populations of human beings amenable to genetically meaningful manipulation by an investigator. As a consequence, in human genetics one is often called upon to deal with collections of observations which are largely unintelligible unless they can be reduced to at most a few numbers, or facts, which summarize the salient features of the data. It is frequently desirable, therefore, to be able to calculate from a set of data a quantity, termed an estimate, which may be taken as the value of a numerically unknown quantity, called a parameter, characteristic of the population from which these data are presumably drawn. To estimate these parameters we assume that the population is distributed in some form, and that one could completely specify the distribution of individuals within this population if the value(s) of the parameter(s) upon which the distribution is functionally dependent was known. Available from which to estimate the value is a set of observations, a sample. Intuitively, we recognize that a given sample may be representative or nonrepresentative of the population. That is, we are aware that if one were to sample repeatedly from the population the samples, and the estimates derived therefrom, would vary one from another. It follows that the only infallible way of ascertaining the value of the parameter involves a study of the entire population, but we presume that this is neither feasible nor possible. It is of importance, then, that the method of estimation which we employ provide estimates whose behavior is predictable. Alternatively stated, there are certain minimum attributes which the estimates should possess. Customarily, we seek estimates which have the properties we term consistency, unbiasedness, and efficiency.

When we assert that a statistic is consistent, we imply that as the sample size increases without limit, the estimate tends to the parameter in question. More formally, an estimate which converges in probability to the estimated value, as the sample size tends to infinity, is called a *consistent* estimate (Cramer, 1946, p. 351). An estimate is said to be *unbiased* if, for fixed sample size, the average value derived from repeated samples converges in probability to the "true value." An estimate can be consistent and yet biased. Thus, for example, as an estimate of the variance, the quantity

$$\sum_{1} \frac{(x_i - \bar{x})^2}{n}$$

where x_i is the observation on the i^{th} individual ($i=1, 2, \ldots, n$) in a sample of size n with mean $\bar{x}$, is consistent but negatively biased; i.e., the mean of a series of such estimates will be smaller than the true variance by an amount which is a function of the sample size. When the bias of an estimate is known, it can, of course, be removed.

We say that an estimate is *efficient* if, for a given and finite sample size, the estimate has minimum variance. In this sense, efficient estimates exist only under rather restrictive conditions, but asymptotically efficient estimates, estimates which possess minimum variance for samples of infinitely increasing size, frequently occur. This concept of efficiency of an estimate is from the late Sir Ronald Fisher who demonstrated many years ago (1922) that a method of estimation, which he termed the method of maximum likelihood, always gives rise to estimates which are asymptotically efficient.

Briefly, the method of maximum likelihood is as follows: Suppose that y is a random variable with a distribution dependent solely upon p. We may indicate this symbolically as f(y;p). If we now assume that a sample of size N is drawn from a population of y's, then the likelihood, say L, of obtaining a given sample will be

$$L=f(y_1; p)\,f(y_2; p)f(y_3; p) \ldots f(y_N; p)=\prod_{i=1}^{N} f(y_i; p).$$

This probability, which is functionally dependent solely upon p, is called the "likelihood function." Now the principle of maximum likelihood states that the best estimate of p is that value which makes the likelihood a maximum. That is, the best estimate of p is afforded by the real positive root, if one exists, of the solution of

$$S(p)\equiv\frac{dL}{dp}=0, \tag{1}$$

or, since L must be positive or at least not negative, the estimate of p may be obtained from the solution of

$$\frac{d \log L}{dp}=0. \tag{2}$$

The latter is generally the more convenient form to use.

Fisher defines the expected amount of information in the sample with respect to p as

$$I(p)=-E\left(\frac{d^2L}{dp^2}\right) \tag{3}$$

or as

$$I(p)=E\left(\frac{d^2 \log L}{dp^2}\right) \tag{4}$$

In large samples we take as the variance of p the inverse of the information; hence

$$\sigma_p^2=(I(p))^{-1}. \tag{5}$$

As an alternative to the latter, one may calculate the variance from the following relationship:

$$\sigma_p^2=\left(\sum \frac{1}{e_i}\left(\frac{de_i}{dp}\right)^2\right)^{-1} \tag{6}$$

where e_i is the expected number in the i^{th} class.

Unfortunately, and all too frequently, particularly when more than one parameter is estimated, the likelihood equation may defy explicit solution. When this occurs, one has recourse to iterative methods of solving the equation, the most common of which has been called maximum likelihood scoring. An especially lucid account of this procedure is to be found in N. T. J. Bailey's *Introduction to the Mathematical Theory of Genetic Linkage.* In brief, one chooses a provisional value for p, say p_1, which makes S(p) nearly zero. One then calculates $I(p_1)$ and obtains an improved estimate

$$p_2=p_1+[S(p_1)/I(p_1)].$$

The score, S(p), and the weight, I(p), are then revised on the assumption that $p=p_2$. This is repeated until the desired accuracy is obtained.

The notions here developed with respect to one parameter can be and have been extended to the case of more than one parameter, but let us now turn to the estimation of gene frequencies.

Gene frequency estimation: If one considers a single diploid locus at which two alleles, A and B, may exist, any one of the following mutually exclusive but exhaustive array of genotype-phenotype correspondences could occur:

1) AA, AB, and BB are all phenotypically distinct;
2) AA is the same as AB, but distinct from BB;
3) BB is the same as AB, but distinct from AA;
4) AA is the same as BB, but distinct from AB;
5) AA, AB, and BB are indistinguishable.

Traditionally, we describe the first of these cases as an instance of "no dominance"; 2 and 3 as instances of "dominance", and for 4 and 5 no common genetic expressions exist. Cotterman (1953) has suggested, however, that case 5, which is both unknown and unknowable, be termed the *identity* system. He further proposed that genotype-phenotype correspondences such as those set out above be called *regular phenotype systems* if all members of a given genotype have the same phenotype, and if this holds for all genotypes. Now some regular phenotype systems are permutationally equivalent. That

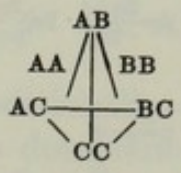
is to say, a change of gene symbols is all that distinguishes one from the other as in cases 2 and 3 above. To identify those systems which were permutationally distinct from the array of all possible phenotype systems, Cotterman coined the expression *phenogram*. It can be readily established, by enumeration, that with 2 alleles, there exist 4 phenograms; whereas with 3 alleles there are 52. If there are more than 3 alleles, enumeration is not feasible, but Bennett (1957) has shown by combinatorial analysis that with 4 alleles there are 5,525 phenograms and with 5, no less than 11,698,156. Each phenogram represents, in effect, a unique problem in gene frequency estimation. The formidableness of an attempt to consider exhaustively the estimation problems associated with 3 alleles, not to say more, is self-evident. However, we know that the gene frequency estimates will be indeterminant if the number of phenotypes, P, does not equal or exceed the number of genes, G, associated with the phenogram. We further know that if P=G, then only one consistent set of estimates will exist; whereas if P>G, more than one consistent set can exist not all of which will be maximally efficient. Thus, in the 2 allele case, at most 3 of the 4 phenograms can give rise to determinate gene frequencies, and, in fact, only 2 do (see table 1). In the case of 3 alleles, at most 42 of the 52 phenograms will yield unique estimates, and some of these must also be indeterminant. The precise number of the latter is not known nor can it be safely conjectured on intuitive grounds. We present in table 2 the 4 phenograms for which estimates exist, and to illustrate the method of maximum likelihood, we derive here a fifth.

Consider a case of three alleles where the genotypes AA, AB, and BB represent one phenotypically distinct group; AC and CC another;

TABLE 1.—*Regular two allele phenotype systems and the appropriate gene frequency estimates. The genotypes which are presumed to be phenotypically indistinguishable are connected by a bar*

Genotype-Phenotype correspondences	Phenotypes	Observed	Expected	Maximum likelihood estimates
AB AA BB	AA AB BB	b c $\frac{d}{n}$	np^2 $2npq$ nq^2	$p=\frac{2b+c}{2n}$ $I=\frac{2n}{pq}$
AB AB or AA BB AA BB	A− BB	a $\frac{d}{n}$	$(1-q^2)n$ nq^2	$p=1-q$ $q=\sqrt{\frac{d}{n}}$ $I=\frac{4n}{1-q^2}$
AB AA——BB	AB AA+BB	c n−c	$2npq$ $(1-2pq)n$	p and q are indeterminate
AB AA——BB	"Identity"			

TABLE 2.—*Some regular three allele phenotype systems and the appropriate gene frequency estimates. The genotypes which are presumed to be phenotypically indistinguishable are connected by a bar*

Genotype-Phenotype correspondences	Phenotypes	Observed	Expected	Maximum likelihood estimates
AB AA BB AC BC CC	AA	a	np^2	$p=\frac{2a+b+f}{2n}$
	AB	b	$2npq$	
	BB	c	nq^2	$q=\frac{b+2c+d}{2n}$
	BC	d	$2nqr$	$r=1-p-q$
	CC	e	nr^2	$V(p)=\frac{p(1-p)}{2n}$
	AC	$\frac{f}{n}$	$2npr$	$V(q)=\frac{q(1-q)}{2n}$
AB AA BB | | AO BO OO	AB	a	$2npq$	$p=p'\left(1+\frac{D}{2}\right)$
	A	b	$(p^2+2pr)n$	$q=q'\left(1+\frac{D}{2}\right)$
	B	c	$(q^2+2qr)n$	$r=\left(r'+\frac{D}{2}\right)\left(1+\frac{D}{2}\right)$
	OO	d	nr^2	$p'=1-\sqrt{\frac{c+d}{n}}$
$V(p)=\frac{p}{8n}\left(4-3p+\frac{pr}{pq+r}\right)$				$q'=1-\sqrt{\frac{b+d}{n}}$
$V(q)=\frac{q}{8n}\left(4-3q+\frac{qr}{pq+r}\right)$				$r'=\sqrt{\frac{d}{n}}$
$\text{Cov}(p,q)=-\frac{qp}{8n}\left(4-\frac{pq}{pq+r}\right)$				$D=1-p'-q'-r'$
AB AA BB | | AC BC CC	AA+AB+AC	a	$(p^2+2pq+2pr)n$	$r=\sqrt{\frac{c}{n}}$
	BB+BC	b	$(q^2+2qr)n$	$p=1-\sqrt{\frac{b+c}{n}}$
	CC	c	nr^2	$q=\sqrt{\frac{b+c}{n}}-\sqrt{\frac{c}{n}}$
$V(p)=\frac{p(2-p)}{4n}$				
$V(q)=\frac{q(2-q(1-p))}{4n}$				
$\text{Cov}(p,q)=-\frac{pq(2-p)}{4n(1-p)}$				
$V(r)=V(p)+V(q)+2\,\text{Cov}(p,q)$				
	AA	b	np^2	$p=\sqrt{\frac{b}{n}}$
	BB	c	nq^2	$q=\sqrt{\frac{c}{n}}$
	rest	a	$(1-p^2-q^2)n$	$r=1-p-q$
$V(p)=\frac{1-p^2}{4n}$				
$V(q)=\frac{1-q^2}{4n}$				
$\text{Cov}(p,q)=-\frac{pq}{4n}$				
$V(r)=V(p)+V(q)+2\,\text{Cov}(p,q)$				

and BC still a third. It should be noted that the gene A is "silent"; it invariably "simulates" B or C. Suppose, now, that in a sample of N, n_1 individuals were of the type (AA or AB or BB), n_2 were of the type (AC or CC), and n_3 were of the type BC. We assume that

this sample was drawn from a population randomly mating with no mutation, migration, nor selective differentials. Thus, we assume that n_1, n_2, and n_3 are multinomially distributed with parameters N, p, and q. The likelihood equation is, therefore,

$$L=\frac{N!}{n_1!n_2!n_3!}(p^2+2pq+q^2)^{n_1}(r^2+2pr)^{n_2}(2qr)^{n_3}.$$

Since $p+q+r=1$, only 2 of the 3 gene frequencies are independent. It follows that if we were to replace r, say, by $1-p-q$, differentiate the resulting expression with respect to p and q, set the partials so obtained equal to zero, and solve the 2 equations simultaneously for p and q we would obtain the maximum likelihood estimates. However, in this case, since P=G, the maximum likelihood estimates may be obtained more simply merely by equating the expected relative frequencies of the 3 phenotypes to their observed relative frequencies. Suppose, then, that $a=n_1/N$, $b=n_2/N$, and $c=n_3/N$; we have

$$p^2+2pq+q^2=a \tag{7}$$
$$r^2+2pr=b \tag{8}$$
$$2qr=c. \tag{9}$$

From (7) we note

$$p^2+2pq+q^2=(p+q)^2=(1-r)^2$$

whence

$$r^2-2r+(1-a)=0$$

and from the quadratic form we have

$$\hat{r}=1\pm\sqrt{a}$$

but only the root less than one can be a gene frequency. If this value for r is substituted in (9), we obtain

$$\hat{q}=\frac{c}{2(1-\sqrt{a})}$$

and, finally,

$$\hat{p}=1-\hat{q}-\hat{r}.$$

Clearly, though A is "silent," its frequency in the population can be estimated.

The variances and covariances of the estimates given above are somewhat more tedious to obtain. However, they may be computed from the information matrix, I, as follows:

$$I=\begin{bmatrix} I_{pp} & I_{pq} \\ I_{pq} & I_{qq} \end{bmatrix}$$

where

$$I_{pp}=2N[2(1-q)^2+q(r+2p)]/r(r+2p)$$

$$I_{pq}=2N[r+pr+2p]/r(r+2p)$$

$$I_{qq}=2N[r^2(2-r)+2pq(1+r)+q^2r]/qr(4+2p).$$

Now the variances are merely

$$\sigma_p^2=I_{qq}/D$$

$$\sigma_{pq}=-I_{pq}/D$$

$$\sigma_q^2=I_{pp}/D$$

where

$$D=I_{pp}I_{qq}-I_{pq}^2.$$

Though the algebraic expressions for the variances and covariances are quite cumbersome, numeric values can be readily obtained in the manner just described.

Once one has generated a set of estimates, what then? Of immediate interest, of course, is the reliability of the estimates; this hinges upon the appropriateness of the genetic model, the randomness of the sample, etc. Let us examine briefly the more important of these. First, is the model appropriate? This can be tested directly from the data from which the estimates are obtained whenever the number of phenotypes exceeds the number of parameters to be estimated by at least one. Under these circumstances one merely compares the observed numbers of individuals in the various phenotypes with those expected, and the significance of the discrepancies which are encountered can be evaluated by chi-square. If, as in the present case, the number of phenotypes does not satisfy the stricture just cited, the appropriateness of the model cannot be tested in the manner described. However, if data exist on two generations, it is generally possible to test the model indirectly through a test based upon the frequency of segregating families (see, for example, Neel and Schull, 1954, pp. 200–203). In either event, it must be borne in mind that the model involves two major assumptions; namely, that the genotype-phenotype correspondences are as postulated and that the population sampled is in Hardy-Weinberg equilibrium. Thus, if the model is deemed inappropriate because of an improbably large chi-square, one does not necessarily conclude that the genotype-phenotype correspondences are not as postulated. The poor fit of the model to the data may result because (a) the genotype-phenotype correspondences are not as postulated but the population satisfies the equilibrium conditions, or (b) the genotype-phenotype correspondences are as postulated but the population is not in equilibrium, or (c) neither assumption holds.

As a general rule, one can distinguish between these alternatives only through further observations.

Second, is the sample random? Of principle concern is the occurrence of correlated observations within the sample. The commonly used estimation equations assume a sample of uncorrelated observations. In much of the published data on gene frequency estimates this restriction is violated. The violations invariably stem from the inclusion of biologically related persons whose genotypes and, therefore, phenotypes are necessarily correlated. We know that the inclusion of relatives tends in general to have little effect upon the estimate itself, but may have a rather marked effect upon the variance of the estimate. If one ignores the relationships within the data, which is tantamount to treating everyone as unrelated, the variance so obtained underestimates the "true" variance. Thus, in a contrast of two populations, one could assert that a significant difference obtains when, in fact, no difference exists. When one is confronted with a sample which includes relatives, three courses of action are open. First, the problem may be ignored as has generally been the case. The possible consequences of this we have just described. Secondly, one may discard the relatives from the sample in such manner that only unrelated individuals remain. In view of the difficulty often attendant upon the collection of the data, this is an unattractive alternative. Finally, one can allow in the estimation procedure for the correlations which occur. The necessary adjustments are not simple. Cotterman (1947) developed a system of weighting appropriate to the two-allele case which gave rise to unbiased, consistent estimates. This method, though simple of use, was not fully efficient. Finney (1948, 1948a) subsequently set out in detail the appropriate maximum likelihood solutions. Most investigators have viewed the computations as too formidable and, as a consequence, relatively little use has been made of the method as judged by the literature. In the case of the ABO blood groups the method proved particularly cumbersome, a deterrrent to its use. Recently, Downs has devised a graphical method which holds promise. Briefly, the procedure is as follows: Suppose one has a sample of size N which is subdivisible into n_1 unrelated individuals, n_2 combinations of parent-offspring, n_3 combinations of siblings, etc. One begins by obtaining an inefficient estimate of p and q, the frequencies of the genes A and B, say. With these estimates one calculates the information matrix associated with n_1, with n_2, etc. Now, the contribution of a single observation (individual) to the matrix of n_1, of a single observation (parent-child) to n_2, etc., can be obtained by interpolation from graphs. Fortunately, the information matrices can be shown to be additive, and to be relatively insensitive to small changes in the values of p and q. Thus, though the estimates of p and q must be obtained iteratively, the information matrices, for practical purposes,

need be computed only once. While some sacrifice in precision follows from the use of graphs and the failure to iterate the variance and covariance estimates, the procedure is so much better than current practices that these small elements of approximation seem unimportant.

Estimation of the coefficient of inbreeding: One of the more frequent departures from random mating, and one of great interest currently is inbreeding. To assert that a population is inbreeding is to imply that relatives marry more frequently than chance would dictate. Such marriages, that is, the marriage of relatives, we term consanguineous, and the offspring derived therefrom we term inbred. Considerable effort has been directed toward delineating the effect of inbreeding on quantitative as well as qualitative variables. No attempt to appreciate these effects could be fruitful without mention of the coefficient of inbreeding. Sewall Wright, to whom the term as well as the concept is due, has defined this coefficient as the "correlation between uniting gametes," the general formula for which he deduced to be

$$F=\sum((\tfrac{1}{2})^{n_s+n_d+1}(1+F_A))$$

"where F_A is the coefficient of inbreeding of any common ancestor that makes the connecting link between a line of ancestry tracing back from the sire and one tracing back from the dam. The numbers of generations from sire and dam to such a common ancestor are designated n_s and n_d, respectively. The contribution of a particular tie between the pedigrees of sire and dam is $(\tfrac{1}{2})^{n_s+n_d+1}(1+F_A)$ and the total coefficient is simply the sum of all such contributions." More recently, Malecot (1948) has defined the coefficient as the probability that an individual will possess at a given genetic locus two genes identical in origin.

An extension of the concept outlined in the previous paragraph, from an individual to a population gives rise to the notion of the average coefficient of inbreeding. This average may be viewed in a number of different contexts. On the one hand, and as the name might suggest, it may be viewed as the probability that an "average" individual will possess two genes identical in origin at a given genetic locus. On the other hand, this average may be viewed as the rate of mixing of two populations, one randomly mating ($F=0$), and one completely inbred ($F=1$). It is the estimation of this average coefficient of inbreeding which is of primary interest to us, and this average may be estimated either directly or indirectly. Let us consider the latter possibility first.

Indirect estimation can proceed in either of two ways. If the population is sufficiently well-known that its breeding formula can be specified, one could obtain an estimate of the coefficient of inbreeding by simulation, *i.e.*, with the aid of digital computers and Monte Carlo

methods. Alternatively, one might attempt to estimate this coefficient much as one estimates gene frequencies. That is to say, one might assume a population randomly inbreeding at some fixed rate, F, wherein the array of genotypes were in a stable equilibrium described by the distribution

$$\sum_i((1-F)q_i^2+Fq_i)A_iA_i+2(1-F)\sum_{i<j}q_iq_jA_iA_j$$

where q_i is the frequency of the gene A_i and the coefficients of the A's describe the frequencies of the various genotypes. If an estimate of F derived from this distribution is to be meaningful, clearly the population must be randomly inbreeding at a fixed rate, and a stable equilibrium must obtain. It is doubtful whether in practice either of these conditions are met; however, the error which is introduced may, in many instances, be negligible.

Li and Horvitz (1953) have shown that from the above distribution a variety of consistent estimates of F can be generated. Among the methods of estimation which they describe are methods based upon the (1) total proportion of heterozygotes, (2) product moment correlation between uniting gametes, (3) determinant of the gametic correlation matrix, (4) value of chi-square assuming panmixia in the population, (5) sum of proportions of alleles in homozygous condition among their respective total frequencies, and (6) method of maximum likelihood. When there are only two alleles, the six methods yield identical expressions for F. This is not so when the number of alleles exceeds two. Unfortunately, the sampling variances to be associated with the first five methods of estimation are not known, and thus it is not clear which, if any, is the method of preference. In the general case, Li and Horvitz were unable to obtain explicit solutions for the gene frequencies and F by the method of maximum likelihood. It is possible, however, to obtain explicit expressions for certain special cases, one of which we wish to consider. We propose to examine a situation such as that which presumably obtains with respect to the ABO blood group system. The argument proceeds as follows:

Consider a genetic model which assumes three genes, no mutation, migration, nor selection, and a population inbreeding at a constant rate. If we let p, q, and r be the frequencies of the three genes (say, I^A, I^B, and I^O) and F the constant rate of inbreeding, then the expected relative frequencies of the four phenotypes are merely,

Phenotype	*Frequency*
O	$rF+(1-F)r^2$
A	$pF+(1-F)(p+2r)p$
B	$qF+(1-F)(q+2r)q$
AB	$2pq(1-F)$.

Now it can be shown that the values of p, q, r, and F which solve the following equations will also solve the maximum likelihood equations:

$$\begin{aligned} Fr+r^2(1-F) &= \overline{O} \\ Fp+(p^2+2pr)(1-F) &= \overline{A} \\ Fq+(q^2+2qr)(1-F) &= \overline{B} \\ 2pq(1-F) &= \overline{AB}. \end{aligned}$$

where $\overline{O}$, $\overline{A}$, $\overline{B}$, and $\overline{AB}$ are the observed relative frequencies of the blood types O, A, B, and AB, respectively. It is not too formidable a task to show from these equations that

$$p=\frac{Z\pm\sqrt{Z^2-8(\overline{AB}+2\overline{B})(2\overline{A}\cdot\overline{AB}+\overline{AB}^2)}}{4(\overline{AB}+2\overline{B})}$$

$$q=\frac{\overline{AB}(1-p)}{(2\overline{A}-p+\overline{AB})}$$

$$r=1-p-q$$

and

$$F=\frac{\overline{0}-r^2}{r-r^2}$$

where

$$Z=4\overline{A}\cdot\overline{AB}+4\overline{A}\cdot\overline{B}+4\overline{B}\cdot\overline{AB}+3\overline{AB}^2.$$

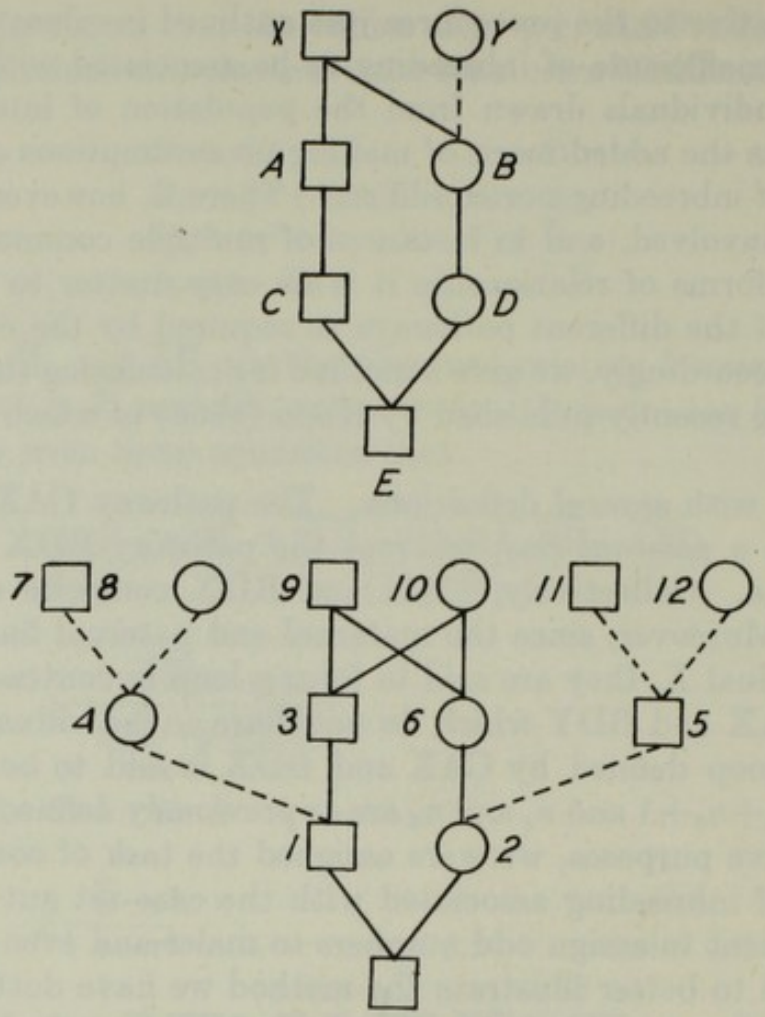
Clearly, there are 2 solutions to the equations, and while the values of p, q, and r will be real in both instances, the values of F may be real or imaginary, positive or negative. Several circumstances can give rise to imaginary values not the least of which is a system markedly out of equilibrium. Among the real values of F, only those which satisfy $1\geq F\geq 0$ are admissible. The variances and covariances of these estimates, which are of a somewhat more complicated form, are to be found in a note published elsewhere (Schull, Ito, and Soni, 1963). For the general case of multiple alleles, it has been argued (see Li and Horvitz, 1953) that since the method of maximum likelihood does not yield the usual gene frequency estimates it may be best to accept the conventional values and estimate F under this set of conditions. The importance of these differing points of view can hardly be overstated. For example, in a sample of 115 individuals of whom 49 are of type A, 35 of type B, 24 of type O, and 7 of type AB, one approach leads to an estimate of F of 0.7781, whereas the other leads to an estimate of approximately 0.00005. A direct estimate of F within this population gives rise to a value of about 0.006 (Schull, Yanase, and Nemoto, 1962). If one accepts this latter estimate as the "true" value, then both methods would appear to be off by a factor of 120 or so, but in different directions.

An alternative to the procedures just outlined involves the calculation of the coefficients of inbreeding to be associated with a random sample of individuals drawn from the population of interest. This approach has the added merit of making no assumptions about either fixed rates of inbreeding nor equilibria. There is, however, somewhat more work involved, and in instances of multiple common ancestors or multiple forms of relationship it is an easy matter to fail to enumerate all of the different pathways as required by the conventional method. Accordingly, we give a method for calculating the coefficient of inbreeding recently published by Kudo (1962) in which errors seem less likely.

We begin with several definitions. The pathway CAX in figure 1 Kudo terms a *paternal line*; whereas the pathway BDX he terms a *maternal line*. Collectively, CAX and BDX comprise an *ancestral line pair*. Moreover, since the maternal and paternal lines share the same individual X, they are said to form a loop in contrast, say, with the lines CAX and BDY which do not share an individual or individuals. The loop defined by CAX and BDX is said to be of length L where $L=n_s+n_d+1$ and n_s and n_d are as previously defined. Suppose, for illustrative purposes, we were assigned the task of computing the coefficient of inbreeding associated with the case set out in figure 2. It is convenient to assign odd numbers to males and even numbers to females, and to better illustrate the method we have dotted the lines of descent of lesser importance. We begin by exhaustively enumerating the paternal and maternal lines. We note (see figure 3) that this enumeration defines 16 ancestral line pairs beginning with the pair 1–3–9 and 2–5–11, and terminating with the pair 1–4–8 and 2–6–10. We further note that as set out in figure 3 each ancestral line pair is represented by a single, unique cell, in an array of 4×4 cells. The Roman numerals serve to identify the generations in the pedigree exclusive of the generation in which the individual in question occurs; the numbers are assigned proceeding backwards in time. It should be noted that the number of paternal or maternal lines will always be 2^{R-1} where R is the maximum generation number when assigned in the manner described. In the present instance, $R=3$, and the number of maternal and paternal lines is 2^2.

We now search systematically for cells corresponding to an ancestral line pair in which a specific individual occurs in both the maternal and the paternal line. For example, consider the first paternal line, namely, 1–3–9. We note that this line shares no common individuals with the first maternal line (2–5–11), nor the second (2–5–12), nor the fourth (2–6–10), but does share an individual, 9, with the third maternal line (2–6–9). We hatch in or in some other manner mark the cell corresponding to the intersection of these lines, that is, lines 1–3–9 and 2–6–9. We continue this search, marking the appropriate cells, until all ancestral line pairs have been examined. In the present case, we

735–712 O—65——5



I II I II III	1 3 9	1 3 10	1 4 7	1 4 8
2 5 11				
2 5 12				
2 6 9				
2 6 10				

find only one additional cell to that described above; the additional cell corresponds to the paternal line, 1–3–10, and the maternal line, 2–6–10. The coefficient of inbreeding we can show is merely

$$F=\frac{\text{number of marked cells}}{2(\text{total number of cells})}.$$

This formula is appropriate to any pedigree with any number of common ancestors provided none of the latter are themselves inbred. Appropriate adjustments can be made, however, to cover this situation (see Kudo, 1962), or the case of sex-linked genes.

Many years ago, Wright and McPhee (1925) proposed an approximate method of calculating coefficients of inbreeding particularly useful in the analysis of unusually complex pedigrees. This method, which consists essentially of a random sampling of ancestral line pairs, would seem to be practicable in those instances, such as the so-called "consanguinity villages" of Japan, where most members of the community are related, and often multiply so. This approach, as well as Kudo's method, is readily adapted to digital computers.

In conclusion: As we stated at the outset, no presentation of the present kind could hope to be exhaustive. Our goals were, therefore, quite limited, namely, to expose some of the methods of population genetics and some of the associated problems. We have attempted to achieve these objectives through a consideration of but two kinds of genetic parameters. Other geneticists would undoubtedly have selected other parameters. Of far greater importance than the choice of the parameter is the methodology, and this would vary little.

Most of us subscribe to the notion that a dialogue between epidemiologists, on the one hand, and geneticists, on the other, must necessarily be fruitful for both our disciplines. A dialogue presupposes, however, that we share some form of communication. Clearly, this is not our specialized jargons which are apt to be mutually incomprehensible, rather it is our common interest in the population study as a means of solving our problems.

REFERENCES

1. Bailey, N. T. J., 1961. *Mathematical Theory of Genetic Linkage.* Oxford: The Clarendon Press.
2. Bennett, J. H., 1957. The enumeration of genotype-phenotype correspondences. Heredity 11: 403–409.
3. Cotterman, C. W., 1947. A weighting system for the estimation of gene frequencies from family records. Contributions of Laboratory of Vertebrate Zoology, University of Michigan, No. 33. pp. 1–21.
4. Cotterman, C. W., 1953. Regular two-allele and three-allele phenotype systems. Amer. J. Hum. Genet. 5: 193–235.
5. Cotterman, C. W., 1954. Estimation of gene frequencies in nonexperimental populations. In: *Statistics and Mathematics in Biology* (Eds.: Kempthorne, O.; Bancroft, T. A.; Gowen, J. W.; and Lush, J. L.) Ames: Iowa State University Press. pp. 449–465.
6. Cramer, H., 1946. *Mathematical Methods of Statistics.* Princeton: Princeton University Press.
7. Crow, J. F., 1963. The concept of genetic load: a reply. Amer. J. Hum. Genet. 15: 310–315.
8. DeGroot, M. H., 1956. Efficiency of gene frequency estimates for the ABO system. Amer. J. Hum. Genet. 8: 39–43.
9. Downs, T. Personal communication.

10. Finney, D. J., 1948. The estimation of gene frequencies from family records. I. Factors without dominance. Heredity 2: 199–218.
11. Finney, D. J., 1948a. The estimation of gene frequencies from family records. II. Factors with dominance. Heredity 2: 369–390.
12. Fisher, R. A., 1922. On the mathematical foundations of theoretical statistics. Phil. Trans. (a) 222: 309–368.
13. Kudo, A., 1962. A method for calculating the inbreeding coefficient. Amer. J. Hum. Genet. 14: 426–432.
14. Levene, H., 1963. Inbred genetic loads and the determination of population structure. Proc. Nat. Acad. Sci. 50: 583–592.
15. Li, C. C., 1963. The way the load ratio works. Amer. J. Hum. Genet. 15: 316–321.
16. Li, C. C. and D. G. Horvitz, 1953. Some methods of estimating the inbreeding coefficient. Amer. J. Hum. Genet. 5: 107–117.
17. Malecot, G., 1948. *Les Mathématiques de l'Hérédité.* Paris: Masson et Cie.
18. Neel, J. V. and W. J. Schull, 1954. *Human Heredity.* Chicago: University of Chicago Press.
19. Sanghvi, L. D., 1963. The concept of genetic load: a critique. Amer. J. Hum. Genet. 15: 298–309.
20. Schull, W. J., Ito, K. P., and A. Soni, 1963. A note on inbreeding and the island of Susak. Jugoslavenska Akademija Znanosti i Umjetnosti (in press).
21. Schull, W. J., Yanase, T., and H. Nemoto, 1962. Kuroshima: The impact of religion on an island's genetic heritage. Human Biology 34: 271–298.
22. Wright, S. and H. C. McPhee, 1925. An approximate method of calculating coefficients of inbreeding and relationship from livestock pedigrees. J. Agri. Res. 31: 377–383.

Hereditary Factors Elucidated by Twin Studies*

By B. Harvald, M.D., and M. Hauge, M.D.,
The Institute for Human Genetics, University of Copenhagen, Copenhagen, Denmark.

INTRODUCTION

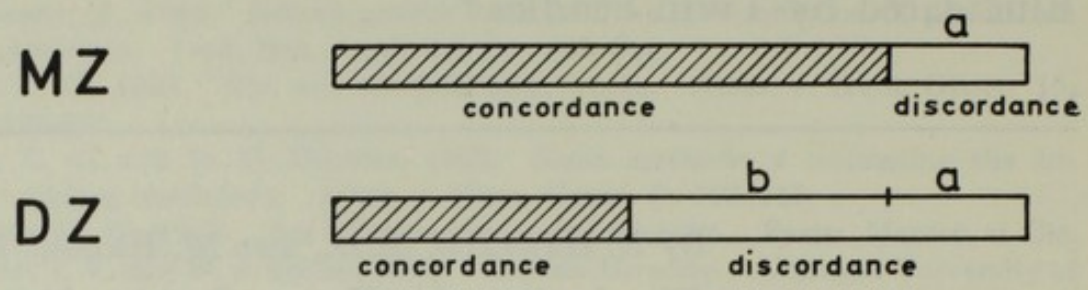
Ever since the days of Galton (1822–1911) twins have been a useful tool in the hands of human geneticists. The contributions of the German psychiatric school have been especially important in furthering the methodology of twin studies. Thus Siemens (1924) elaborated a polysymptomatic similarity test for the diagnosis of zygosity and Luxenburger (1940) laid down rules for correct sampling techniques.

The principle of the classical twin method is simple enough and illustrated in table 1. Monozygotic twins (MZ) carry identical genes while dizygotic (DZ) or fraternal twins are genetically as different as ordinary brothers and sisters. Phenotypical differences between MZ partners must be due to environmental differences (a), whereas DZ partners may also differ because of different genetic endowments (b). On the assumption that the intra-pair difference of environment is the same for MZ and DZ twins and therefore will cause a phenotypical difference (a) of the same magnitude in the two groups, it may be concluded that a trait which shows a substantial difference in the degree of resemblance encountered in pairs in the two groups must, to a certain extent, depend on genetic factors.

The comparison may be made on the basis of all-or-none data as well as of measurement data. In the former case, twin pairs where both partners show the same trait are called concordant, whereas pairs where only one of the partners is affected are called discordant. A significantly higher concordance rate in MZ than in DZ pairs, for instance with regard to a special disease, indicates a genetic influence. A concordance rate approaching 1.00 in the MZ twins and at the same

*This investigation was supported by research grant C-948 and GM 09418-02 from National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare, Bethesda, Md.

THE PRINCIPLE OF THE CLASSICAL TWIN METHOD.



a = difference due to environment

b = difference due to gene effect

For all - or - none data:

$$(1) \quad H = \frac{Conc._{MZ} - Conc._{DZ}}{1 - Conc._{DZ}}$$

For quantitative traits:

$$(2) \quad H = \frac{var._{DZ} - var._{MZ}}{var._{DZ}}$$

time significantly higher than in DZ pairs will be found in instances where genetic factors are of paramount etiological importance.

Various formulas have been elaborated in order to reach a precise expression of the relative influence of genes and environment with regard to the variation of different traits in human populations. One of these is the formula of the H-value (Holzinger, 1929) given in table 1 which attempts to utilize concordance rates in MZ and DZ twin pairs to derive a quantitative estimate of heritability. This formula is subject to many criticisms. First, the value of H depends almost totally on the definition of concordance with regard to the trait in

question. For diabetes mellitus, concordance may mean the manifestation of diabetes in the co-twin before a given age, the manifestation at any age, and finally the finding of an abnormal glucose tolerance test without manifest diabetes. In epilepsy, concordance may mean chronic epilepsy in the co-twin, rare or a single epileptic seizure during the whole life, or seizures only after such special provocation as fever, infection, or head injuries. The concordance rates in these cases will differ substantially resulting in very different values of H.

The estimate of heritability on the basis of measurement data as shown in formula (2) is somewhat more meaningful, but here too some reserve with regard to the interpretation is called for. First of all, the assumption that the intra-pair difference of environment is on an average the same in MZ and DZ pairs is dubious. For monochorionic MZ pairs the difference of the intra-uterine environment in some cases may be extreme because of a common placenta, in consequence of which one of the twins may "starve out" the co-twin by virtue of vascular anastomoses. This factor may cause an underestimate of H. During childhood parents are inclined to stress the similarity of MZ twins and therefore treat them more alike than is the case in DZ pairs, where the twins themselves are generally very eager to stress their dissimilarities. Neglect of such biases will attribute too large a share to heredity as a factor in the diversity of non-identical twins. A factor working in the same direction is the tendency of MZ twins to stick together, "hunt together" as the psychologists like to call it. Thus, because of their special twin situation, they subject themselves to the same environmental influences.

Of quite another nature is the fact that MZ twins, because of their identical genetic complement, will actively choose environments which may show striking intra-pair similarity, such as similar occupations and similar partners. This is a fine example of the intimate interaction of genes and environment, and it may be considered a question of taste, if the influence of this similarity of environment after all should not be taken as an expression of a gene effect.

Reflections of the same sort may be applied to the relationship between DZ same-sexed twins (DZ_1) and DZ different-sexed twins (DZ_2). If the concordance rate in DZ_1 pairs exceeds that of DZ_2 pairs this implies a sexual dependance of the manifestation of the trait in question, either because of endogenous factors or because of environmental factors differing between males and females. Generally the DZ_2 pairs are excluded and the H-value calculated on the basis of the MZ and DZ_1 pairs only, but whether this procedure is correct is debatable.

Special caution is called for when dealing with small samples or unreliable tests. The significance level of a given value of H is difficult to assess, but still the H-value has the merit of simplicity in spite of all limitations.

TYPES OF TWIN STUDIES

The different ways in which twin studies may be approached are shown in table 2.

TABLE 2.—*Different types of twin studies*

Twin case reports	Concordant MZ pairs. Discordant MZ pairs.
Twin series	Comparison between MZ and DZ pairs. Comparison between MZ twins brought up apart and MZ twins brought up together.
"Twin control method"	Comparison between MZ twinpartners one of whom has been under specific environmental influence.

Case reports, usually of *concordant* MZ pairs, crowd the literature of human genetics. Even if the value of such single cases is limited, as they can neither prove the presence of a gene effect nor allow any estimate of the relative importance of genetic factors, case reports both on common and rare traits have often been useful in indicating a field for further investigation, either as family studies or based on the collection of twin series.

Detailed case reports of *discordant* MZ pairs are by far more conclusive and may be utilized in the analysis of environmental differences. Especially in psychiatry, this method has proven useful in elucidating the etiologic importance of different exogenous factors with regard to the trait in question.

Comparison of MZ and DZ pairs is the classical twin method discussed above, being informative with regard to the nature/nurture relationship, as well as the question of "false heredity," so often left open by family studies, which cannot decide whether an intrafamilial accumulation of cases is due to common genes or common environment. On the other hand, only on rare occasions can these data be used to test some specific genetic hypothesis. The method provides no answer to the problem whether dominant or recessive genes, single-gene or polygenic combinations, or one or many genotypes are responsible for the appearance of a trait.

Further drawbacks of this method are the difficulty of getting a sufficient number of affected twins in the case of uncommon traits, and the difficulty of representative sampling. This means the twin approach will usually be limited to common traits. On the other hand, the method is suitable for the study of quantitative traits and traits with a complicated mode of inheritance, where family studies are often inconclusive. The sampling technique is difficult and there will be a tendency toward overrepresentation of concordant pairs. A detailed discussion of these questions has been given by G. Allen (1955).

Comparison of MZ twins brought up apart and MZ twins brought up together suggests whether similarities in MZ twins brought up together may be more or less due to similarities of their environment,

in this way serving as a sort of control of the results of the ordinary twin method. Furthermore, twins brought up apart allow an estimate of the extent to which environment may change a given character. The drawbacks of the method are, on the one hand, the extreme difficulty of getting a sufficient amount of material because of the rarity of such twins, and, on the other hand, the fact that the foster homes in which these twins are found generally do not differ as much as we would wish them to. The method has thus far been restricted to the study of different anatomical, physiological, and psychological traits (Newman *et al.*, 1937, Juel-Nielson, Mogensen, 1957, and Shields, 1962).

The designation, "the twin control method," has been used for the study of MZ twin pairs, where one of the partners has been subjected to some specific environmental influence, and where a comparison of the affected and the unaffected partners may be informative. A study of this sort has been carried out with twins who had sustained a head injury (Dencker, 1958) and who later on complained of sequelae of the injury. It could be demonstrated that the symptoms which the head-injured patients ascribed to their injury were found with approximately the same frequency in the co-twins, who had not sustained any head injury. Thus it might be concluded that the symptoms had very little or nothing to do with the head injury, but could be ascribed to the premorbid psychic constitution of the patient. In pharmacotherapeutic assays, where the drug is given only to one of the partners of MZ pairs, the twin control method makes it possible to draw conclusions from materials of very limited size.

Whatever method is used, it is of primary importance to make a correct *zygosity diagnosis*. Generally the zygosity diagnosis is quite simple. Especially in monochorionic MZ twins, however, there will be a tendency to classify the more dissimilar pairs as DZ. Because the series nearly always are of a limited size, a wrong classification of even a single or a few pairs may highly influence the results. The extensive use of blood and serotyping has, in this connection, proved indispensable. In this way it is possible to obtain a group of twins who have a 100 percent probability of being dizygous, and another group with a high probability of monozygosity.

THE TWIN STUDY OF THE INSTITUTE FOR HUMAN GENETICS

As mentioned above the problem of representative sampling has been an obstacle in many earlier twin studies. This point is clearly demonstrated by an overrepresentation of MZ pairs compared with DZ pairs and of concordant pairs proportional to discordant pairs. In order to avoid both of these factors, in 1954 we planned an extensive

followup study of all twins born in Denmark in a definite period, *i.e.*, the years 1870–1910. This program has now very nearly been accomplished, and, in the following, some of our results and conclusions will be discussed with the purpose of illustrating the practical and theoretical difficulties met with in this sort of study.

In the period 1870–1910, 37,914 twin births (including stillbirths) were recorded in Denmark. Because of the very high infant mortality of earlier years, it may be estimated that between 50 and 60 percent of the pairs born in the period mentioned were broken by the death of one or both partners before the age of 5 years. These early broken pairs have been excluded for several reasons, namely, 1) the difficulty of establishing a reliable zygosity diagnosis in the same-sexed pairs, 2) the very incomplete medical data which it has been possible to procure in these cases, and 3) the fact that the whole project has been focused upon the study of pathological traits appearing in the later years of life.

Including the losses by emigration this leaves an estimate of less than 15,000 pairs available for study. Until now satisfactory contact has been established with 6,893 pairs, where full details of the medical history have been obtained. A further 2,000 pairs have been traced but the information is still insufficient. This, however, leaves 5–6,000 pairs which we have not been able to trace. This may imply an uncontrollable bias, which it has not been possible to avoid.

The diagnosis of zygosity was based on a modified similarity test. A questionnaire was sent to all twins of identical sex, containing questions about the degree of similarity between the partners of the pair. Pairs were considered probably monozygous (MZ) when the information obtained either from the twins themselves or from close relatives familiar with both of them indicated that they were conspicuously similar in appearance and that no difference in eye or hair color was found. If, further, reliable information was presented to the effect that a twin had repeatedly been confused with his partner, this has been taken as a very reliable indication of a high probability of monozygosity. When the answers to the questions about conspicuous similarity between the partners were clearly in the negative, the pair was placed in the group of probably dizygous twins (DZ_1). Wherever vague, uncertain, or conflicting answers were obtained and the zygosity diagnosis was therefore doubtful, extensive blood group examinations were used where both partners of a pair were alive at the time of examination. Where this was not the case, the pair was classified as being of uncertain zygosity.

In a small sample of the material, comprising 165 same-sexed pairs, we have checked the method used in diagnosing the zygosity by subjecting these pairs to a thorough examination of their blood and serum groups including the following systems: A_1A_2BO, MNS, Rh (with 5

antisera), P, Lewis, Duffy, Kell, Lutheran, Gm, Hp, and Gc. It may be calculated that complete identity between the twin partners as regards all these tests gives on the average an MZ probability of about 98 percent. Out of 78 pairs which had been classified as MZ according to the questionnaire, all except 1 pair presented no difference with regard to blood or serum group. Eighty pairs considered DZ according to the questionnaire showed one or more intrapair blood group differences in 76 cases, whereas the remaining 4 pairs had identical blood groups. Out of the total of 165 pairs, 7 had been classified as being of uncertain zygosity on the basis of the questionnaires. Four of these had identical blood and serum groups. Thus it turns out that our diagnostic procedure is sufficiently reliable.

Table 3 shows the composition of that part of the material which has been included in the present analysis. It appears that the material agrees with the Weinberg (1901) formulation, according to which the number of the MZ twins is the total number minus 2 times the number of different-sexed pairs. The cases of uncertain zygosity diagnosis are probably mostly MZ pairs, which makes an MZ/DZ ratio of very nearly 1/3, which is the rate of the average Danish population.

TABLE 3.—*Material included in the present analysis*

Same sex:	
Probably MZ	1,528
Probably DZ_1	2,609
Type of zygosity uncertain	231
Different sex (DZ_2)	2,525
Total	6,893

Most of the pairs for which the diagnosis of zygosity is uncertain are pairs in which both partners have died before the time of examination. This, too, implies a source of bias.

The information concerning the medical history of the twins has been procured first by way of questionnaires sent to all living twins, or, in case of death of both partners, to some surviving near-relative of the twins. Information was requested about all admissions to hospitals, whatever the cause had been, and about symptoms of any of a number of specified diseases. If the questionnaires remained unanswered, the twins or their relatives were visited by one of the investigators.

All stays in hospitals have been verified by a review of records. Supplementary information has, if necessary, been obtained from general practitioners. All death certificates and autopsy records of deceased twins have been checked irrespective of the suggested cause of death.

As might be calculated from the total number of twins, an adequate number of pairs could be obtained only with regard to common diseases. A diagnosis of *cancer* was made in a total of 1,038 individuals. In a number of these cases, however, the cancer diagnosis had not been verified either histologically, by operation or by autopsy. Tables 4 and 5 show the concordance rate in different zygosity groups. All individuals in whom a diagnosis of cancer was established were considered index cases. Therefore in the calculation of the concordance rate both partners of the concordant pairs are probands, and the pair will be counted twice.

TABLE 4.—*Malignant growth in twins (1038 pairs)*

Concordance rates with regard to cancer of same site			
	MZ	DZ_1	DZ_2
Verified cancer	8/160	8/335	4/321
Cancer diagnosis uncertain	6/47	4/62	4/49
Total	14/207	12/397	8/370
		$0.05 > P > 0.01$	
Zygosity diagnosis uncertain			4/64

TABLE 5.—*Malignant growth in twins (1038 pairs)*

Concordance rates with regard to cancer of all sites			
	MZ	DZ_1	DZ_2
Verified cancer	17/160	43/335	36/321
Cancer diagnosis uncertain	16/47	12/62	8/49
Total	33/207	55/397	44/370
Zygosity diagnosis uncertain			12/64

It is necessary to distinguish two different degrees of concordance: full concordance with regard to cancer of the same localization as the cancer of the index case, and concordance with regard to cancer of all localizations. If all pairs where either the cancer diagnosis or the zygosity diagnosis is not well established are left out of consideration, no significant differences are found between MZ and DZ pairs. Thus in spite of the considerable size of the material it cannot be proved on this basis that hereditary factors are at all active in cancer.

The group with uncertain cancer diagnosis is characterized by a rather high number of concordant pairs, especially with regard to cancer of the same localization. This probably reflects the inclination of physicians to suspect cancer, and especially cancer of the same localization, as a cause of death in a patient known by him to

have had a co-twin who had died from cancer. Thus the concordance rates found after inclusion of these doubtful cases are no doubt far too high, and the comparison between the concordance rates of MZ and DZ_1 pairs must be looked upon with scepticism.

When dealing with a disease showing a relatively late age of onset and a prolonged period of manifestation, such as is the case in cancer, some sort of age correction is called for. This may be done in different ways. According to Weinberg (1927) discordant pairs for which the nonaffected co-twin has not yet entered the manifestation period are not counted at all. Pairs where the co-twin has not passed the whole period are counted half. This procedure is not, however, very practical in cancer, where a period of manifestation is difficult to define. The same objection may be raised to the more elaborate procedure of Strömgren (1935), where each pair is to be counted exactly according to the fraction of the manifestation period which has been passed by the nonaffected co-twin. Slater (1953) has suggested calculating the probability of the healthy co-twin having manifested the disease at the time of examination on the basis of the differences between the ages of manifestation in concordant twins. This method, too, is rather time consuming and requires a rather considerable number of concordant pairs for a meaningful calculation of probabilities.

In our material we have tried a more simple age correction, by leaving out of the analysis all pairs where the healthy co-twin has died before the age of manifestation of malignant growth in the proband. Even after this, however, the concordance rate with regard to cancer of all localizations remains virtually unchanged in all zygosity groups, viz., 0.15 for MZ, 0.18 for DZ_1, and 0.14 for DZ_2.

Our figures support the conclusion that genetic factors determining susceptibility to cancer in general do not play any detectable role in human cancer. The existence of genetic factors determining the development of malignant growth in specific organs cannot be totally ruled out, but on the basis of the rather low concordance rates it may be estimated that their importance, compared with unknown exogenous factors, is very small.

Cancer exemplifies the application of the twin method in the case of a rather well-defined all-or-none character, where the diagnosis may be made with reasonable certainty. In many diseases, however, the calculation of concordance rates is not too meaningful. This, for instance, is the case in different forms of arteriosclerotic disease, where it would be more meaningful to try to get a quantitative measure of the degree of arteriosclerosis in the co-twins. An investigation of this sort will be difficult and has not been practicable in our material. Instead, we have chosen to use the occurrence or nonoccurrence of *coronary occlusion* as an indicator of the degree of arteriosclerosis. As is apparent from table 6 no differences as to the concordance rate

can be demonstrated between MZ and DZ_1 pairs. The difference at the 5 percent significance level between DZ_1 and DZ_2 on the other hand indicates the unequal sex-distribution in coronary occlusion, which could be due either to gene effects or different environmental influences in males and females.

TABLE 6.—*Cardio-vascular diseases in twins*

	MZ	DZ_1	DZ_2
Coronary occlusion	20/102	24/155	10/133
(390 pairs)			0.05>P>0.01
Arterial hypertension	20/80	10/106	4/106
(292 pairs)	0.01>P>0.001		
Cerebral apoplexy	22/98	16/148	22/179
(425 pairs)	0.05>P>0.01		

In *arterial hypertension*, on the contrary, a difference significant at the one percent level may be demonstrated between MZ and DZ_1 pairs. The absolute figures are of limited value, as they will only indicate concordance with regard to clinically established hypertensive disease. An estimate of the relative genetic influence in determining the blood pressure necessitates a determination of the intra-pair differences of blood pressure in MZ and DZ pairs in different age groups and at different blood pressure levels. A program of this sort is going on at present, but the material is still too small for any conclusions. Also in *cerebral apoplexy* a difference between MZ and DZ_1 pairs can be demonstrated.

With regard to *psychiatric disorders* the material is not very suitable. An intensive psychiatric study is not possible as so many of the probands and their co-twins have died or their original psychiatric pattern has been changed in old age. Only a few comments, therefore, shall be added to the survey given in table 7. The total number of institutionalized *oligophrenics* is only slightly higher than the number expected based on the total number of institutionalized oligophrenics in Denmark, indicating that after the fifth year of life, mental deficiency is not more common in twins than in the general population. In *epilepsy* all types have been included, as the inhomogeneous clinical examinations of the probands do not allow a distinction between genuine and symptomatic epilepsy. The total number of epileptic pairs is about double as high as expected from the general population. This may either be a reflection of the higher number of birth injuries in twins, or indicate that the estimate for the general population is too low.

A demonstration of genetic factors influencing resistance to *infectious diseases* and *antibody formation* is given in table 8. Whereas no significant difference could be demonstrated between the concordance rates in MZ and DZ_1 pairs with regard to *death from acute infections*, a highly significant difference is found in *tuberculosis*. It

cannot be taken for granted, however, that this difference is at all due to genetic factors. In the case of infectious diseases the tendency of MZ twins to "hunt together" may influence the results, partly because they will more often be exposed to the same contacts and partly because of their higher possibility of infecting each other.

TABLE 7.—*Psychiatric disorders in twins*

	MZ	DZ_1	DZ_2
Mental deficiency (67 pairs)	12/18 0.001>P	0/20	0/29
Epilepsy (127 pairs)	10/27 0.05>P>	6/43 0.01	4/57
Manic-depressive psychosis (55 pairs)	10/15 0.001>P	2/23	0/17
Schizophrenia (71 pairs)	4/9	4/33	2/29
Suicide (96 pairs)	4/21	0/38	0/37

TABLE 8.—*Infections and antibody production in twins*

	MZ	DZ_1	DZ_2
Death from acute infection (581 pairs)	10/127	14/221	26/233
Tuberculosis (648 pairs)	50/135 0.001>P	42/267	36/246
Rheumatic fever (576 pairs)	30/148 0.001>P	12/226	14/202
Rheumatoid arthritis (188 pairs)	16/47 0.001>P	2/71	8/70
Bronchial asthma (256 pairs)	30/64 0.01>P>	24/101 0.001	22/91

In *rheumatic fever* and *rheumatoid arthritis* it may similarly be doubted, if the difference between MZ and DZ_1 twins indicates an hereditary tendency to these specific forms of antigen-antibody reaction. The higher concordance rate in MZ pairs might equally well be the result of the more intimate contact of MZ partners with common streptococcal infections and common environmental influences. With the same reservations, the rather high concordance rate in MZ twins with *bronchial asthma* may be interpreted as a reflection of the hereditary character of allergic reactions.

Caution is called for in the evaluation of these different concordance rates, based as they are solely on the clinical manifestation of the trait in question. This may be clearly demonstrated in peptic ulcer, where a geographically delimited part of the material with both partners alive has been carefully scrutinized with X-ray examination of both the probands and their co-twins (Gotlieb Jensen, 1964). Table 9 shows that the concordance rates found in this way are considerably higher, but still the MZ–DZ_1 difference is unimpressive,

as an indication that the intrafamilial accumulation of ulcer cases must be primarily due to mutual environmental factors of ulcerogenic effect, and not to mutual genes.

TABLE 9.—*Peptic ulcer in twins*

	MZ	DZ_1	DZ_2
Total material, clinical data only	30/144	24/209	18/243
(596 pairs)	0.208 0.05>P>	0.114 0.01	0.074
Intensive study:			
After exclusion of early broken pairs	8/27 0.296	14/62 0.226 0.01>P>	2/52 0.038 0.001
After X-ray examination	10/21 0.476	18/40 0.450	7/30 0.233

Table 10 shows the findings in *diabetes mellitus* and serves to illustrate the importance of some sort of age correction in cases where a more exact estimate of the concordance rate is desired. In the total material a highly significant difference is found between the rate of 0.47 in MZ and of 0.10 in DZ_1. After exclusion of all discordant pairs where the co-twin has not been observed for at least 70 years, the rates have changed to 0.73 and 0.34, respectively. Calculation of an H-value in the first case gives 0.43, in the second case, 0.59. Age correction (Slater, 1953) gives a concordance rate of 0.71 in MZ and 0.22 in DZ_1 with an H-value of 0.63. The application of Slater's method, however, is complicated by considerable uncertainty in establishing the exact time of manifestation of the disease in the instances where only one twin is affected. It therefore was found necessary to make use of the time when the diagnosis was made instead of the time of manifestation. However, several extraneous factors may influence the time at which the diagnosis is made, especially in elderly persons, where the diabetes is often diagnosed by chance in connection with a medical examination for quite other reasons.

In the material of Then-Bergh (1939) where pairs, one of whom had clinical diabetes and the other of whom had a diabetic glucose-

TABLE 10.—*Diabetes mellitus in twins*

	MZ	DZ_1	DZ_2
Total material	36/76	12/122	10/116
(314 pairs)	0.47	0.10 0.001>P	0.09
After exclusion of pairs where the healthy co-twin has not reached the age of:			
30 years	0.49	0.11	0.09
50 years	0.55	0.13	0.09
70 years	0.73	0.34	0.29
Age correction a.m. Slater	36/51 0.71	12/55 0.22	10/55 0.18

tolerance-test were considered concordant, the concordance rate in MZ twins approached 1.00, also making the H-value approach 1.00. This illustrates the difficulty of deriving a meaningful H-value for clinical traits where the manifestation period is not well-defined and the expressivity of the gene is varying.

In several clinical family studies a genetic association between different clinical traits has been suspected. In the families of patients with Graves' disease an accumulation of cases of other types of thyroid disease has been demonstrated (Bartels, 1941). Similarly, a significantly increased incidence of cancer of the stomach has been found in the families of patients with pernicious anemia (Mosbech, 1953). In the present twin material, with the complete medical histories of more than 13,000 people, there is an opportunity of further elucidation of problems of this sort. Table 11 gives a survey of the occurrence of different diseases in the nondiabetic co-twins of our diabetic probands. A considerable number of these persons is expected to possess the diabetes gene or genes. In diseases occurring with another frequency in this group than in the general population, it must be considered if this may somehow be due to a pleiotropic effect of the diabetic genotype. In most of the examples chosen there is no significant difference between the number expected and the number observed (for example, with respect to gallstones, coronary occlusion, and tuberculosis: disorders very frequent in the diabetic patients themselves). This may indicate that these disorders are a consequence of the diabetic disease, and not a direct gene effect. On the other hand, violent death and death from surgical complications are significantly more common among the nondiabetic co-twins. We have no satisfactory explanation of this finding. Malignant growths are less frequent in the nondiabetic co-twins than expected, a probable sign of some sort of protection from cancer connected with the diabetes genotype.

TABLE 11.—*Occurrence of different diseases in 256 nondiabetic co-twins of diabetic probands*

	Expected	Observed
Peptic ulcer	11	13
Gallstone	10	10
Tuberculosis	12	13
Coronary occlusion	9	11
Rheumatic fever	11	9
Arterial hypertension	6	9
Cerebral apoplexy	8	12
Bronchial asthma	5	9
Malignant growth	23	13
	0.05>P>0.01	
Death from acute infection	11	12
Violent death	6	17
	0.01>P	
Death from complicated surgery	5	16
	0.001>P	

735-712 O—65——6

The only measurement data which until now have been procured from the material with reasonable certainty are the *ages at death*. As shown in table 12 the intrapair difference in pairs, where both partners are dead at the time of examination, is significantly smaller in MZ than in DZ_1 and DZ_2 pairs. On the basis of the intrapair variances an H-value may be calculated, which is supposed to indicate the fraction of the variance in DZ_1 twins due to genetic factors. For the total material the value is 0.29.

TABLE 12.—*Longevity in twins*

	MZ	DZ_1	DZ_2
Average intrapair difference of age at death (years)	14.5±0.94	18.6±0.84 0.01>P>0.001	20.4±0.99
Intrapair variance	209	293	378
$\left(\frac{\Sigma \text{ diff.}^2}{2\times\text{number of pairs}}\right)$			

$$H=\frac{\text{var.}_{DZ_1}-\text{var.}_{MZ}}{\text{var.}_{DZ1}}=0.29$$

The lifespan of the individual may be considered the final result of the cooperation of all gene effects and all environmental influences during life. The H-value calculated on the basis of the lifespan variances in the present material has been found to be considerably higher in the urban than in the rural population. This may be due to a more uniform environment in the urban districts, and suggests that the progressive urbanization of society may gradually lead to a state where the variation of longevity throughout the population is predominantly determined by genetic factors.

The absolute figures must be considered with reserve, first of all because the pairs broken before the fifth year of life have not been included in the calculation. Moreover, restriction of the calculation to pairs where both partners have died by the time of examination may mean a bias of unpredictable effect.

Several other items are being intensively studied in our twin material. A criminologic research program has been started; the results, however, are not ready for publication as yet. The same is the case in our study of fertility in twins, which seems to show a significantly smaller difference in MZ than in DZ pairs with regard to the number of offspring.

This paper has aimed at illustrating the numerous difficulties met within a clinical twin study, among which the difficulty of getting a material of sufficient size should be stressed. Even in a material of the present size only the very common disorders may be elucidated. The analysis is complicated by diagnostic uncertainty, which in combination with sampling biases may lead to fallacies of different kinds. In a way the negative results, as ours with regard to cancer,

are more conclusive than the positive ones, where the justification of a comparison between different zygosity groups in nearly all cases may be doubted.

SUMMARY

The basic principles of twin studies are discussed. So are the limitations of the classical twin method where concordances in MZ and DZ pairs brought up together are compared. Doubt whether environmental differences for identical twin partners are really of the same magnitude as for nonidentical twins is considered the most important objection to the validity of the conclusions drawn from this method. The problems met within clinical twin studies are illustrated from a Danish twin material with followup of nearly 7,000 pairs of twins born in the years 1870–1910. From this material it may be concluded that genetic factors play an unimportant role in the etiology of malignant growths, coronary occlusion, and peptic ulcer. On the other hand, a significantly higher concordance rate in MZ than in DZ pairs could be demonstrated in arterial hypertension, cerebral apoplexy, mental deficiency, epilepsy, manic-depressive psychosis, tuberculosis, rheumatic fever, rheumatoid arthritis, bronchial asthma, and diabetes. Furthermore it can be demonstrated that the intrapair difference of age at death is significantly smaller in MZ than in DZ twins.

REFERENCES

1. Allen, G., 1955. Comments on the analysis of twin samples. *Acta Genet. Med. et Gemel.* 4: 143–160.
2. Bartels, E. D., 1941. *Heredity in Graves' Disease.* Munksgård, Copenhagen.
3. Dencker, S. J., 1958. A followup study of 128 closed head injuries in twins using co-twins as controls. *Acta psychiat. et neurol.* 33: Suppl. 123.
4. Gotlieb Jensen, K., 1964. Peptic ulcer in twins. To be published.
5. Holzinger, K. J., 1929. The relative effect of nature and nurture influences on twin differences. *J. Educat. Psychol.* 20: 241–248.
6. Juel-Nielsen, N. and Mogensen, A., 1957. Uniovular twins brought up apart. *Acta Genet. Med. et Gemel.* 7: 430–433.
7. Luxenburger, H., 1940. Zwillingforschung als Methode der Erbforschung des Menschen. In: *Handbuch der Erbbiologie des Menschen* by G. Just, Volume II, 213–248. Springer, Berlin.
8. Mosbech, J., 1953. *Heredity in Pernicious Anaemia.* Munksgård, Copenhagen.
9. Newman, H. H., Freeman, F. H., and Holzinger, K. J. 1937. *Twins: A Study of Heredity and Environment.* University of Chicago Press, Chicago.
10. Shields, J., 1962. *Monozygotic Twins, Brought Up Apart and Brought Up Together.* Oxford University Press, London.
11. Siemens, H., 1924. *Die Zwillingspathologie. Ihre Bedeuting, Ihre Methodik, Ihre Bisherige Ergebnisse.* Springer, Berlin.
12. Slater, E., 1953. Psychotic and neurotic illnesses in twins. *Medical Research Council Special Report Series.* 278: 1–385.

13. Strömgren, E., 1935. Zum Ersatz des Weinbergschen "abgekünzten Verfahrens". Zugleich ein Beitrag zur Frage von der Erblichkeit des Erkrankungsalters bei der Schizophrenie. *Ztschr. f. d. ges. Neurol. u. Psychiat.* 153: 784–797.
14. Then-Bergh, H., 1939. Zur Frage der psychischen und neurologischen Erscheinungen bei Diabeteskranken und deren Verwandten. *Ztschr. f. d. ges. Neurol. u. Psychiat.* 165: 278–283.
15. Weinberg, W., 1908. Über den Nachweis der Vererbung beim Menschen. *Jahresh. Verein f. vaterl. Naturk. in Württemberg.* 64: 368–382.
16. Weinberg, W., 1927. Mathematische Grundlagen der Probandenmethode. *Zeitschr. Abstgs. und Vererbgsl.* 48: 179–228.

ABO Blood Groups, Secretor Status, and Susceptibility to Chronic Diseases: An Example of a Genetic Basis for Family Predispositions

By J. A. Fraser Roberts, M.D., D. Sc., F.R.C.P., *Clinical Genetics Research Unit, Medical Research Council, Institute of Child Health, The Hospital for Sick Children, Great Ormond Street, London.*

The search for possible associations between blood groups and disease has a long history, in fact it goes back for rather more than 40 years. The earlier literature was, however, confused, negative and conflicting. There were several reasons for this, of which the most important was that the numbers studied were far too small. In retrospect it can be seen that there were two positive results. Buchanan and Higley in 1921 did produce results which foreshadowed the association between group O and peptic ulceration and between group A and pernicious anaemia. But their own conclusion was that no association existed. Years later Ugelli (1936) in Italy got the same result for ulcer. These isolated suggestive findings made little or no impact, however, for they were swamped in the mass of inconclusive results. It was not until 1953 that the first study to carry more or less general conviction was published (Aird *et al.*, 1953). This study was based on very large numbers and showed with very high probability that cancer of the stomach is about 20 percent more frequent in persons of group A than in those belonging to group O.

Soon afterwards an even stronger association was found between group O and peptic ulceration (Aird *et al.*, 1954). This time the numbers were really large and the significance of the association so high that I think most people were convinced that an empirical or

epidemiological finding had been confirmed with high probability. Since then a very considerable literature has been produced. Some associations seem established beyond reasonable doubt, others seem highly probable, while with others there are no more than suggestive indications, or indeed a plain conflict of evidence. And, of course, it is clear that the great majority of diseases show no blood group associations or at least nothing measurable with practicable numbers.

I have said that the great majority of observers and critics seem to be convinced that some associations at least are proved empirically beyond reasonable doubt. One further point may be mentioned in passing. That such associations would be discovered had been predicted, notably by Fisher and Ford. The theoretical importance of the associations is, in fact, considerable, with all the bearings they have on the conflict between those who uphold the neutral, or nearly neutral gene, and those who uphold balanced polymorphism.

A minority of critics, however, notably Dr. A. S. Weiner (1962), are not even convinced that the empirical evidence is sound; and though perhaps it is a small minority it is articulate, so I think I ought to present the evidence for just one disease. This is duodenal ulceration, for which the evidence is most overwhelming; it is also the disease which I propose to talk about almost exclusively in this paper. Once we are convinced that at least one single association is reasonably proved, other associations for which the evidence is more problematical can be left for the clarification that time will bring.

Table 1 shows an up-to-date compilation by Hanley (1963) of the results for duodenal ulceration in persons of group O compared to group A. The third column shows relative incidences. These are obtained by simple cross-multiplication of the figures in the disease and corresponding control series. A relative incidence of 1.59, as shown in the first line, means that the incidence of ulceration in people of group O is 1.59 as compared with unity in people of group A. The total numbers are now very large, more than 16,000 people with duodenal ulcer. The series come from centres all over the world, from Great Britain, the United States, Denmark, Norway, Austria, France, Italy, Greece, Japan, India, West Africa, Israel and Australia. Hanley has compared the relative incidence in people of group O as against those of group A. A comparison of group O versus the other three groups, A, B and AB, added together would have given even lower probabilities. It will be seen that in all 22 series the incidence in group O is higher than in group A. At no less than 16 centres individually the difference is highly significant.

The total χ^2 can be divided into two parts. The first, with one degree of freedom, measures the divergence of the estimated mean value of 1.38 from the 1 expected if there were no association. It is no less than 297, or 17 times its standard error. This corresponds

TABLE 1.—*The analysis of the incidence of duodenal ulcer in persons of group O relative to its incidence in persons of group A (from Hanley)*

Centre	Number in disease series	Relative incidence O:A	χ^2
London	946	1.59	38.82
Manchester	423	1.27	4.89
Newcastle	482	1.46	14.06
Liverpool	1489	1.33	11.05
Glasgow	1642	1.10	2.37
Copenhagen	680	1.42	17.36
Oslo	579	1.49	19.62
Vienna (1)	1160	1.21	7.74
Vienna (2)	449	1.42	10.89
Iowa (whites)	2386	1.47	71.37
Peking	211	1.59	6.89
London	564	1.28	7.19
Tokyo	163	1.67	7.09
Ibadan	390	1.30	3.86
Jerusalem	158	1.46	3.81
Middlesbrough	722	1.48	21.49
Paris	1195	1.40	25.02
Bombay	380	1.16	1.10
Perugia	628	1.52	21.38
Copenhagen	1223	1.33	19.74
Athens	147	1.04	0.02
Sydney	192	2.22	21.50
Total	16,209		337.26

Mean weighted relative incidence=1.38.
Ninty-five percent fiducial limits=1.33 to 1.43.
Difference from unity (degrees of freedom=1) $\chi^2=297.48$.
Heterogeneity between centres (degrees of freedom=21) $\chi^2=39.78$.

to an infinitesimal probability. It is far smaller than a billionth of a millimeter in the orbit of Neptune, or a billionth of a milligram in the estimated mass of the universe. The fiducial limits of the estimate are quite narrow; there is a 95 percent probability that the true value lies within the bracket 1.33 to 1.43.

The remainder of the total χ^2 represents the heterogeneity of the 22 centres; it is 40 for 21 degrees of freedom. There is significant heterogeneity, the probability being about 1 in 2,000. But I think the point to note is how small it is compared to the divergence of the total averaged incidence from unity. Anyway, it would be surprising if the relative incidence were to be the same at such varied centres in so many different parts of the world, where the incidence of the disease and the relevant environmental factors are so different.

From the beginning of these observations in their modern phase it was pointed out that although the associations might be real enough they might be secondary, and so of little real importance. It might be that there are strains in the populations who are, say, high in group O and at the same time especially prone to duodenal ulceration, just as examination of people in an eastern European city might well show a secondary association between group B and dark hair, which would merely be a reflection of racial origin. The variety of the series, however, from so many different parts of the world, and their general concordance, makes this explanation very unlikely. Nevertheless, a direct test is desirable. Such a test was proposed by Pro-

fessor Penrose. It can be seen in sibships which are segregating for ABO blood groups whether the index case with the ulcer is unduly often of group O. In other words, we can see whether people of group O are more likely to develop duodenal ulceration than their own non-O sibs. If they are, the hypothesis of stratification is disproved. The labour of carrying out this test is truly Herculean, but it has been done, thanks to the pertinacity of Clarke and his associates in England and Buckwalter and his associates in the United States. Their combined results are again quoted by Hanley (1963). His figures include recent data supplied by Clarke *et al.* and hitherto unpublished.

TABLE 2.—*Analysis of the change of the propositus being group O in the sibships segregating for blood group O. (Method of analysis that of C.A.B. Smith—see Clarke* et al., *1956)*

(From Hanley, 1963)

Centre	Total sibships	Group O propositi		Variance
		Observed	Expected	
Liverpool (Clarke *et al.*, 1963)	159	86	78.0927	37.3136
Iowa (Buckwalter *et al.*, 1960)	96	56	46.9797	22.0082
Total	255	142	125.0724	59.3218

The difference divided by its standard error ($\sqrt{\text{variance}}$)

$$=\frac{16.9276}{\sqrt{59.3218}}=2.198. \quad p=<0.05.$$

The combined results show that in segregating sibships the observed number of index cases who are of group O is 142, as compared with 125 if there were no association. The difference is significant at the 5 percent level, which is all that can be expected with these small numbers. The difference is, in fact, fully consistent with a primary association and significantly against the hypothesis of an association secondary to stratification.

Personally I fail to see how anyone can feel doubtful about the reality of the association between blood group O and duodenal ulceration. But, of course, the evidence is empirical, or epidemiological, in the broad sense. It is the same kind of evidence as that for an association between cigarette smoking and lung cancer. No causal mechanism has yet been discovered either with duodenal ulceration or with any of the other conditions which show blood group associations. Hazarding a guess, I feel inclined to agree with those who suggest that the mechanism is likely to turn out to be immunological.

You are, of course, aware of the recent literature on possible associations between ABO blood groups and infective diseases. None of these can be regarded as proved as yet, but there are the suggestions about group A and smallpox and group B and plague. There is also

the finding of McDonald and Zuckerman (1962) of a highly significant excess of group O amongst those with upper respiratory illness who show evidence of infection with the Asian influenza virus, A2; this work, however, needs confirmation. Then there is the curious finding resurrected from the literature of the twenties that people of group O became more often, and with high significance, seronegative after anti-syphilitic treatment than did people of group A. It may indeed be that clues as to the meaning of the associations will emerge first from studies on micro-organisms. But I must return to the subject of chronic diseases.

Another association, which I think we must accept as proved with very high probability indeed is that between nonsecretion and duodenal ulceration. Secretion or, alternatively, nonsecretion in the saliva and some other body fluids of substances having ABO specificity depends on a single gene pair, secretion being dominant to nonsecretion. The loci for the ABO genes and for the secretor-nonsecretor genes are not genetically linked. Actually the association between nonsecretion and duodenal ulceration is closer than that between group O and ulceration, but as the numbers are so much smaller the probabilities are not so infinitesimal. Nevertheless, the evidence is overwhelming.

As Doll *et al.* (1961) have shown recently, the separate effects of ABO blood group and of secretion must be regarded as multiplicative rather than additive. It should be mentioned again here that persons of groups A, B and AB go together in respect of susceptibility to duodenal ulceration. There is no difference in relative incidence between persons of these three groups; and all differ by the same amount from persons of group O. Hence we can think of two blood group classes only, namely O and not-O. Thus there are four phenotypes to consider: secretor, not-O; secretor, O; nonsecretor, not-O; nonsecretor, O. These can be symbolized SN, SO, sN, sO. The relative incidence of duodenal ulceration in the four phenotypes can be calculated. The results are shown in table 3.

TABLE 3.—*Relative incidence of duodenal ulceration in subjects of the four phenotypes*

Phenotype	Relative Incidence	
	Doll	Clarke
Secretor, not-O (SN)	1	1
Secretor, O (SO)	1.4	1.3
Nonsecretor, not-O (sN)	1.8	1.7
Nonsecretor, O (sO)	2.7	2.4

One column shows figures taken from the results of Doll *et al.* (1961), the other the results from Liverpool (including additional unpublished figures) as given by Hanley (1963). I should mention

that with Doll's results I have combined three series without weighting to give a single series of figures. It will be seen how close these estimates are. Taking the lowest incidence, that in persons who are secretors and not-O, as unity, the relative incidence rises steadily, until in the limit, with people who are nonsecretors and group O, it is no less than 2.4 or 2.7.

I think that there is little doubt that there is a genetic element in the etiology of duodenal ulceration. The best study I know is that of Doll and Buch (1950). They found that 45 out of 562 brothers of men with ulcer were similarly affected. The expected number, based on a comparable control population and matched age, was 16.67. Thus the incidence amongst brothers was 8.0 percent, against an expectation of 3.0 percent, or 2.7 times as great. Doll and Buch concluded that this resemblance between brothers must be largely genetic. One reason for this was that the resemblance between fathers and sons was much the same, and, of course, brothers share an environment which is more closely similar than do father and son.

ABO blood groups and secretor status are wholly determined by heredity, and relatives tend to resemble each other in these respects to a greater degree than do unrelated people. It is possible, therefore, to calculate how much of the resemblance between brothers in respect of duodenal ulceration can be attributed to their resemblance, for genetic reasons, in ABO blood group and secretor status. In other words, we can estimate how much of the resemblance in susceptibility to ulcer is due to the effect of genes at two known independent loci.

TABLE 4.—*Comparison of incidence of ulcer in general population and in brothers*

1	2	3	4	5	6
Phenotype	Population frequency	Column 2 times Doll's relative incidence of table 3	Expected frequencies of those with D.U.	Expected frequencies in brothers	Column 5 times Doll's relative incidence of table 3
SN	.3825	.3825	.262	.326	.326
SO	.3675	.5145	.353	.368	.515
sN	.1275	.2295	.158	.143	.257
sO	.1225	.3308	.227	.163	.441
Total		1.457			1.539

In the calculations of table 4, I have used the figures of Doll *et al.* (1961), as given in table 3. These are the relative incidences: secretor, not-O, 1; secretor, O, 1.4; nonsecretor, not-O, 1.8; nonsecretor, O, 2.7. We also need figures for the gene frequencies. For simplicity I have taken the frequency of O to be .7; of not-O, .3; of secretion, .5; of nonsecretion, .5. These are in fact close to the real frequencies in Great Britain and in much of North America. They are just about those that would be observed in northern England.

Given these estimates, the frequencies in the population of persons belonging to the four phenotypes are given in column 2. We can now multiply by the factors of table 3, which gives us column 3. Reducing the total to unity, we get column 4, which shows the expected frequencies of the four phenotypes amongst those suffering from duodenal ulceration. Actually, these estimates correspond closely to the observed frequencies amongst those with ulcers. Applying the Hardy-Weinberg law, which works so well in blood group calculations, we can estimate the frequencies of the four phenotypes amongst brothers of those with duodenal ulcers. The result is given in column 5. The brothers are naturally intermediate in blood groups and secretor status between the index cases and the general population. Finally, just as before, we multiply these last figures by the factors of table 3. This gives column 6, which gives a measure of the amount of duodenal ulceration in brothers. The ratio of the total of column 3 to the total of column 6 gives the relative incidence of ulcer in brothers to that in the general population, as determined, that is, by the genetic resemblance of brothers in regard to ABO group and secretion-nonsecretion. The answer comes to 1.056. That is, the effect of genes at these two loci makes the incidence of duodenal ulcer in brothers of index cases 5.6 percent higher than in the general population. Taking Doll and Buch's estimate of an increase in frequency of ulceration in brothers of 170 percent over that in the general population, we see that only 3 percent of the degree of resemblance between brothers is to be ascribed to the known genes at two loci which are known to affect susceptibility, and this in spite of the considerable effect they have on degree of susceptibility in terms of the individual.

What are we to make of this finding? It was clear on the face of it that resemblance in susceptibility to a disease and the side effect, if one may use that loose term, of a single major gene must be small, unless indeed the difference in susceptibility of the different phenotypes is very great indeed, and this is a very unlikely thing to find. What it does fit in with is the usual concept of multifactorial inheritance as the basis for genetically determined differences in resistance and susceptibility to common diseases. The genes at the two known loci are in this particular respect forming part of a multifactorial system, and the small contribution they make reinforces the conception of multifactorial systems based on the additive effect of many genes. It fits in with Darlington and Mather's (1949) conception of polygenic inheritance, a criterion for which is that the effect of any single gene pair shall be small in relation to the total variation.

There is nothing odd about the finding that ancillary effects of major genes play their part in multifactorial systems. This has never been denied by those who have been the foremost advocates of multifactorial determination, though they would think it likely

that the bulk of multifactorial determination is due to genes of small effect, genes which do not have a major effect in some other context. The idea is in fact at least 45 years old. I wish I had been able to verify my reference here, but I have not; it is all so long ago. Perhaps Professor Neel, or someone else who has been a real *Drosophila* geneticist, can help me. [*Editorial note:* cf. Dobzhansky, T., 1927. Studies on the manifold effect of certain genes in *Drosophila melanogaster*. Zirt. f. ind. Abstam.—und Vererbungslehre 43: 330–388.] My recollection is that somewhere about 1916, one of the earlier geneticists—it may have been Sturtevant—measured the size of certain organs, for example, the length of spermathecae, in wild type and white-eyed flies. He found that the gene difference made small differences to these metrical characters.

It seems to be generally true that in animals, as contrasted to plants, differences in genetically determined degrees of resistance and susceptibility to infective diseases are multifactorially determined. The same thing is probably equally true of genetic differences in respect of chronic and noninfective diseases. In fact I would go a step further. Evidence seems to be accumulating that even with defects and deformities, whose causation is partly genetic, the genetic basis may also be multifactorial, at least in some instances. One very important finding here is that of my colleague, Dr. Cedric Carter, on congenital pyloric stenosis (Carter 1961a, 1961b). The sex incidence in this condition is very unequal: one boy in 200 male live births; one girl in every 1,000 female births. What Carter has found is that the incidence in sibs and children of affected women is notably higher than in sibs and children of affected men. For example, about 7 or 8 percent of sons or brothers of affected men are affected, but no less than about 20 percent of sons and brothers of affected women. This is quite incompatible with any single gene hypothesis. Women who are affected are affected in spite of the protection conferred by their sex; they are persons of unusually high genetic susceptibility, and this is faithfully mirrored in their relatives. There must be, in fact, degrees of susceptibility—that is, a continuous distribution and, pretty certainly, a multifactorial basis.

We are finding indications of the same thing with harelip. (Roberts, Carter and Buck, in preparation.) There are two conflicting pieces of evidence. In the first place the incidence in sibs and in children of affected persons is the same. This points to a dominant gene, of partial penetrance of course. In the second place, when two or more relatives other than sibs are affected, they tend quite often to be on both sides of the family, which equally clearly points to a recessive gene. Our feeling is that the evidence really points to neither. On the other hand, multifactorial determination would fit perfectly well and we have some preliminary indication of the same sex relationship as in congenital pyloric stenosis. I can also mention some results on

talipes equino-varus, which go the same way (Wynne-Davies, 1963).

But I have strayed too far from chronic diseases, and will only say that the general character of inheritance seems likely to be multifactorial. There is also the point that the genetic element is usually relatively small, though not invariably so, in relation to the nongenetic factors involved in causation. But, of course, even if small it is nonetheless important in understanding the etiology of the various conditions.

A final word might be added on blood group and disease associations in general. The proposition that they do not exist must, I think, be rejected. The next objection, that they might be secondary to racial or some other stratification seems exceedingly unlikely, to say the least. Furthermore, in the one instance in which the hypothesis has been put to the test it has been contradicted by the findings. But some might still say: so what? What do these associations matter in practical terms? Of course the evidence remains epidemiological. No satisfactory explanation, no indication of mechanisms, have yet emerged. But I feel sure they will. Epidemiology is, of course, the beginning and not the end of research. It shows that there is a problem to investigate and may point the way to useful lines to follow. If due note had been taken of the fact that at the Retreat, a mental hospital at York founded by Quakers and mostly taking in Quaker patients, general paralysis was almost unknown at a time when it represented a substantial fraction of admissions to ordinary mental hospitals, then the connection between syphilis and general paralysis might have been discovered far earlier than it actually was.

There have been some interesting findings in regard to blood group and disease associations, but nothing very meaningful so far. I can add one new observation, due to Hanley (1963) whose summaries I have already quoted. Using an improved technique for measuring levels of serum pepsinogen, he was able to show that persons of group O have on the average higher levels than persons of group A. This is in normal subjects, not those with ulcer. The difference is very highly significant. There is also a sex difference. Interestingly, there is no difference between secretors and nonsecretors, so presumably a different mechanism may be involved. I do feel sure, however, that in due time the underlying mechanisms will be found and that the result will be a useful contribution to the study of several diseases.

REFERENCES

1. Aird, I., Bentall, H. H., and Roberts, J. A. Fraser, 1953. Relationship between cancer of the stomach and ABO blood groups. Brit. Med. J. *1:* 799–801.
2. Aird, I., Bentall, H. H., Mehigan, J. A., and Roberts, J. A. Fraser, 1954. Blood groups in relation to peptic ulceration and carcinoma of the colon, rectum, breast, and bronchus; association between ABO groups and peptic ulceration. Brit. Med. J. *2:* 315–321.

3. Buchanan, J. A., and Higley, E. T., 1921. The relation of blood groups to disease. Brit. J. Exp. Path. *2:* 247–255.
4. Carter, C. O., 1961a. The inheritance of congenital pyloric stenosis. Brit. Med. Bull. *17:* 251–254.
5. Carter, C. O., 1961b. *In* Clinical Aspects of Genetics, edited by F. Avery Jones. London: Pitman.
6. Darlington, C. D., and Mather, K., 1949. The Elements of Genetics. London: Allan and Unwin.
7. Doll, R., and Buch, J., 1950. Hereditary factors in peptic ulcer. Ann. Eugen. *15:* 135–146.
8. Doll, R., Drane, H., and Newell, A. C., 1961. Secretion of blood group substances in duodenal, gastric and stomal ulcer, gastric carcinoma, and diabetes mellitus. Gut *2:* 352–359.
9. Hanley, W. B., 1963. Hereditary aspects of duodenal ulcer. M.D. Thesis, University of Liverpool.
10. McDonald, J. C., and Zuckerman, A. J., 1962. ABO blood groups and acute respiratory virus disease. Brit. Med. J. *2:* 89–90.
11. Ugelli, L., 1936. Distribuzione dei gruppi sanguingni negli individui portatori di ulcere gastroduodenali. Policlinico (sez prat.) *43:* 1591–1594.
12. Weiner, A. S., 1962. Blood Groups and Disease (A Critical Review). Lancet *1:* 813–816.
13. Wynne-Davies, R., 1963. Personal communication.

Epidemiological Concepts Applied to Studies of Chronic Diseases*

By Abraham M. Lilienfeld, M.D., *Department of Chronic Diseases, Johns Hopkins School of Hygiene and Public Health, Baltimore, Md.*

Students of the chronic diseases have become increasingly interested in the distribution of disease in the population and of those factors that influence this distribution, that is, in the epidemiological aspects of disease. The value of such knowledge has been repeatedly demonstrated in the study of the infectious diseases and of some diseases of current medical and public interest such as heart disease and various forms of cancer. There is an increasing indication that such knowledge is useful to the geneticist who is concerned with the population aspects of his discipline. This afternoon, time permits only a rather brief summary of this subject. Since many of the aspects have been reviewed recently in various articles and textbooks, we would first like to provide a broad overall view of epidemiological methods in general terms and, then, illustrate the application of some of these methods to two forms of cancer in which we have been interested.

In studying the distribution of a disease in a population, attention is paid to variations in disease frequency by such factors as time, place and certain individual characteristics of those affected by the disease. Thus, the epidemiologist studies the changes in the frequency of the disease over a period of time, its distribution by various geographic areas and the frequency of disease by age, race, sex, socioeconomic class, individual living habits such as smoking, and exposure to agents such as radiation. He is interested in determining what characterizes the diseased person in contrast to the nondiseased in the population.

*Some of the studies, whose results are reported in this paper, were supported in part by a Public Health Service career program award number GM-K6-13,901 from the National Institute of General Medical Sciences, by a grant from the Mary Reynolds Babcock Foundation, Winston-Salem, North Carolina, and by Training Grant No. CT-5085 from the National Cancer Institute.

Basically, the epidemiologist attempts to determine whether there are statistical associations between a disease and certain population characteristics or attributes of possible etiological importance. Thus, the first phase of the epidemiological approach to a disease problem is the determination of these statistical associations. The second phase is the derivation of biological inferences from the pattern of statistical associations. It is the totality of the series of statistical associations and these biologic inferences that are termed the "epidemiology of the disease."

The methods of determining statistical associations can be classified as follows:

(A) *Studies of General Population Characteristics*
  1. Vital Statistics including mortality and morbidity data obtained on a routine basis.
  2. Special Population Surveys.

(B) *Studies of Individual Characteristics*
  1. Observational Studies
    (a) Cross-sectional study
    (b) Retrospective study
    (c) Prospective study
  2. Human Experimental Studies

In our discussion of specific examples, we will illustrate several of these methods. Their relative advantages and disadvantages have been discussed in previous publications (Lilienfeld, 1959; MacMahon *et al.*, 1960).

Once these associations have been determined, the two general types of hypotheses that must be considered to explain these associations are: (1) a causal hypothesis, which states that the factor associated with the disease is a cause of the disease, or (2) an indirect association hypothesis, which states that the statistical association is indirect and noncausal and that it is a reflection of the existence of a common underlying X factor which predisposes the individual to have the disease as well as to have the characteristic. The problem then consists of determining the relative plausibility of these two hypotheses. This is usually done by integrating data obtained from experimentation, both animal and human, studies of biochemical and pathogenetic mechanisms and various types of further epidemiological studies. Some of the issues involved in these problems have been recently discussed (Yerushalmy and Palmer, 1959).

With this structural outline as a background, we would like to review some of the epidemiological aspects of breast cancer to illustrate the types of data and the types of reasoning used by the epidemiologist.

We have already indicated that the epidemiologist is interested in the age distribution of a disease. Figure 1 contains the average annual age-specific death rates for breast cancer by sex. These rates are plotted on semi-logarithmic paper, which represents the

rate of change of mortality with age. The death rates for females suggest that over a lifetime, they can be resolved into two linear components, one consisting of those prior to and one of those after the 40 to 45 year age group and that the rates in the second part of the curve have a lower slope than those in the first. This observation has been interpreted as being consistent with the hypothesis that the proportion of women in the general population who are "susceptible" to breast cancer decreases after 40 to 45 years of age. Since this coincides with the average age of onset of menopause, these data suggest that entrance into menopause decreases susceptibility to breast cancer. With respect to the males there is also a suggestion that the rate of increase of mortality, as shown by the change in the slope of the line is slightly less after 40 to 50 years of age.

Another possible explanation for this change in slope which must be considered is that we are dealing with two separate sets of etiological factors. DeWaard *et al.* (1960) suggested that the underlying etiological factor for breast cancer was excessive estrogenic activity, but that prior to 40 to 45 years of age, ovarian estrogenic activity was of etiological importance, whereas after the menopause it was the estrogenic activity of adrenal origin. They attempted to look at this indirectly and presented some data to suggest that more of the postmenopausal breast cancer patients had obesity and hypertension than a control group, whereas this was not observed in the premenopausal group of cancer patients. Their study leaves a great deal to be desired but further studies comparing the characteristics of these two groups of breast cancer patients with controls would appear worthwhile. Differences in these two groups with respect to a history of nursing habits have been reported by Kamoi (1960). He found that female breast cancer patients over 45 years of age gave a history of nursing their children for shorter periods of time than a control group whereas those between 35–44 years of age did not show such a difference. We shall discuss the pertinence of length of nursing to breast cancer later.

The most consistently observed factors associated with female breast cancer concern marital and pregnancy habits. Breast cancer is more frequent among those never married than among those who marry and breast cancer patients on the whole marry at a later age than do women in general. The latter association results in their age at first pregnancy being later than usual and possibly also in their having less children, an observation which has been consistently made (Lilienfeld, 1961; Wynder *et al.*, 1960).

In figure 2 there are presented the age-specific incidence rates of reported cases of breast cancer in 1949 among single and married women in upstate New York; the term "single" applies to women reported as such at the time of report in contrast to those who were reported as being married, widowed or divorced, which have been

735–712 O—65——7

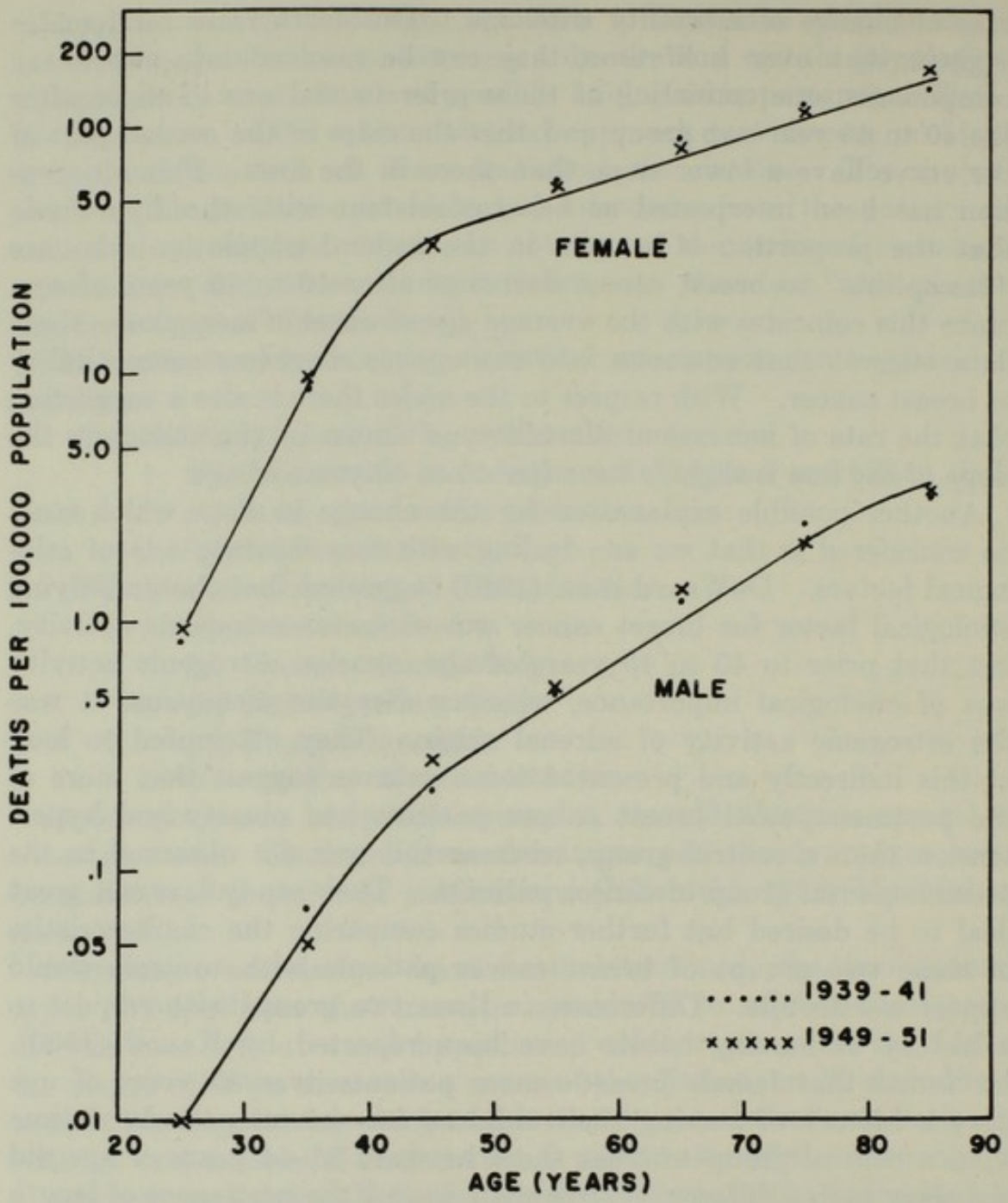


FIGURE 1.—Average annual age-specific death rates (per 100,000 population) for breast cancer by sex, white population, U.S., 1939–41, 1949–51.

grouped together as the "married" group. It is apparent that the resolution of the incidence rates into two linear components is present for both the single and married groups. The rates comprising the first component are almost identical for both the single and the married women. The excess incidence of female breast cancer among single women appears to be limited to the latter period of life, after about 40 years of age. Similar patterns have been observed in selected metropolitan areas of the United States, the Netherlands, and England and Wales (Dorn and Cutler, 1955; Stocks, 1955; Versluys, 1955).

In accordance with our previous interpretation of the age pattern, one possible explanation for the variation of incidence with marital

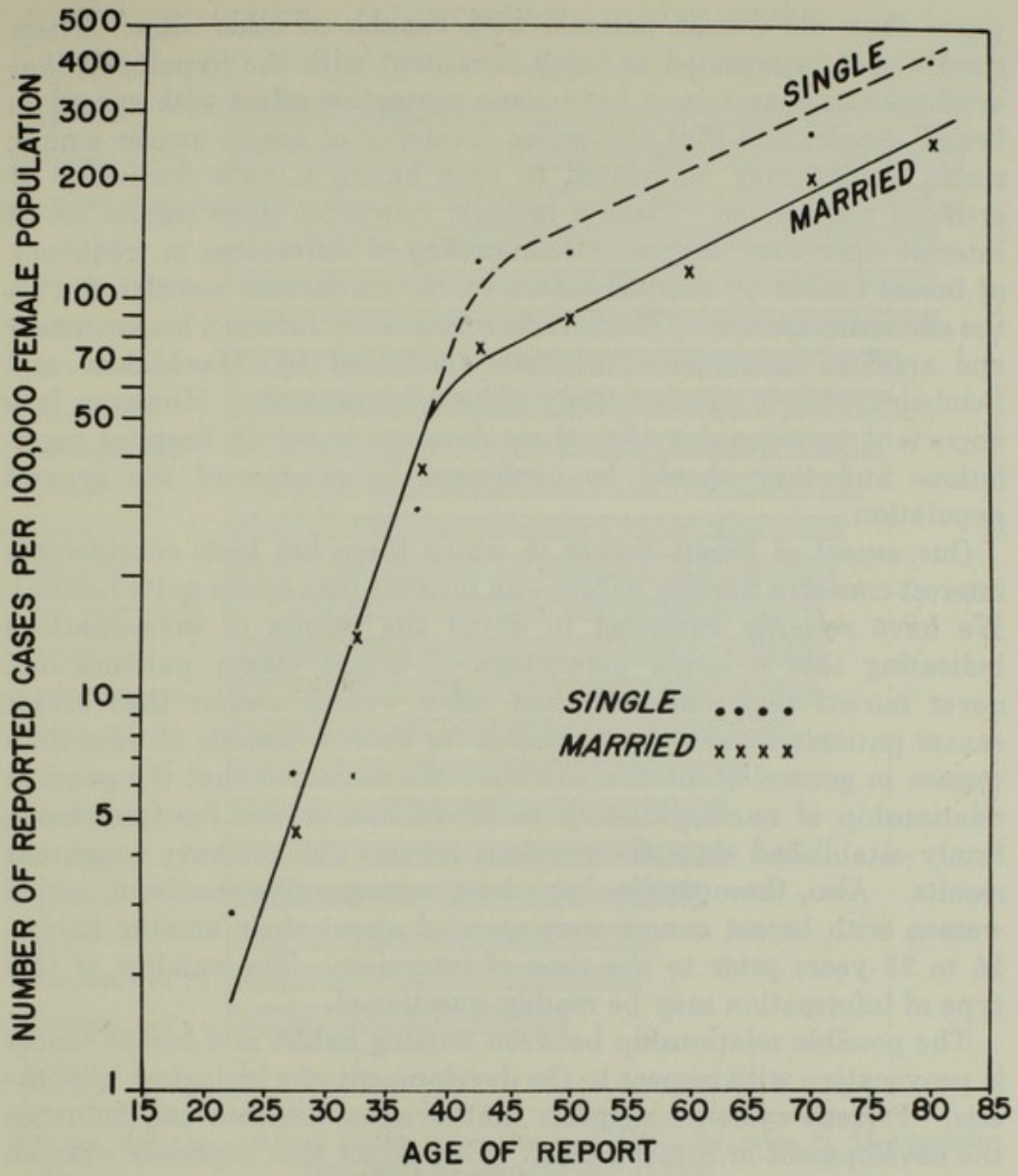


FIGURE 2.—Annual age-specific incidence rates of reported cases of cancer of the female breast by marital status, New York State, exclusive of New York City, 1949.

status may be that single women as a group enter the menopausal state later, and consequently remain in the more highly susceptible state longer than do married women. In an attempt to test this hypothesis, the menopausal histories of patients admitted to the Roswell Park Memorial Institute were studied; the results have already been reported in detail (Lilienfeld, 1956). Briefly, we found that the frequency of artificial menopause was 30 percent less among single than among married women and that artificial menopause occurred about 10 years earlier than natural menopause. No difference in time of occurrence of natural menopause in the two groups was noted. Also, patients with breast cancer had a lower frequency of artificial meno-

pause than did female patients with cancers of other sites. These results were interpreted as being consistent with the hypothesis that artificial menopause may have some protective effect with regard to breast cancer, and that the excess incidence of breast cancer among single women may be related to their having a lower frequency of artificial menopause. From a biologic viewpoint these results are of interest since they suggest a relationship of differences in frequency of breast cancer by marital status to certain factors associated with the endocrine system. The inverse relationship between breast cancer and artificial menopause has been confirmed by MacMahon and Feinleib (1960) in another study of hospital patients. However, it is important to remember that these data are based on hospital populations and they should be confirmed by studies of the general population.

One aspect of breast cancer in which there has been considerable interest concerns nursing habits—an interest that seems quite natural. We have recently reviewed in detail the results of investigations indicating that a larger percentage of breast cancer patients had never nursed their children than other women and/or that breast cancer patients nursed their children for shorter periods of time than women in general (Lilienfeld, 1963). We indicated that the possible relationship of nursing history to breast cancer was far from being firmly established since the previous reports did not have consistent results. Also, these studies have been retrospective studies in which women with breast cancer were queried about their nursing habits, 15 to 25 years prior to the time of interview. The validity of this type of information may be readily questioned.

The possible relationship between nursing habits and breast cancer is provocative with respect to the development of a biological hypothesis. Present evidence suggests that ovarian function can influence the development of breast cancer. One factor that depresses ovarian function is prolonged nursing. Thus, nursing may not have a direct protective effect on breast cancer but an indirect one through its depressive influence on ovarian function. The hypothesis is attractive since it has been invoked as a possible explanation for the differences in mortality rates of female breast cancer in different countries.

To illustrate this, figure 3 contains a comparison of the age-adjusted mortality rates from female breast cancer in several countries for the period 1952–1954 as collected by Segi *et al.* (1959). In comparing these rates one must keep in mind the differences in death certification practices as well as the availability of medical care in these countries. The most striking fact emerging from this comparison is the remarkably low death rate in Japan. Wynder *et al.* (1960) have suggested that it may be a result of the fact that Japanese mothers nurse their children for a two-year period, which is longer than in the other countries.

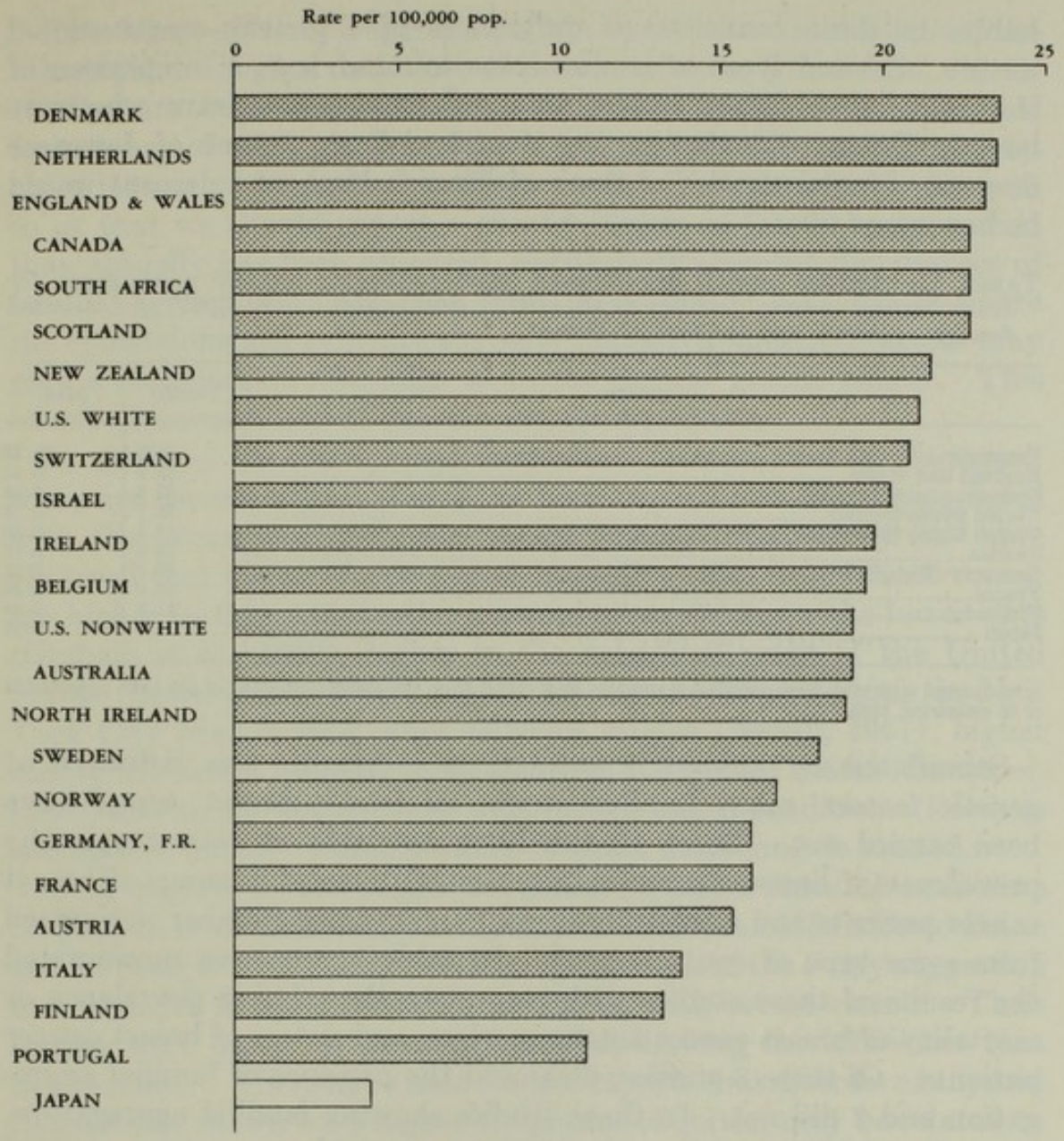


FIGURE 3.—Age-adjusted female breast cancer mortality rates in 24 countries, 1952–54, according to Segi (1959).

We thought it would be of interest to make a similar comparison with respect to death rates of breast cancer among males; these are presented in table 1, which indicates that for *both* males and females the death rates are lower in both Finland and Japan. The fact that the low mortality rates in Japan are present for both sexes raises questions as to whether differences in breast feeding habits are a reasonable explanation for differences in the frequency of breast cancer among females. If one assumes that these two diseases are similar (which may be an erroneous assumption), it would appear that there are factors present in Finland and Japan which decrease the risk of *both* the male and female inhabitants developing and dying from breast cancer. These differences in mortality should be studied further, particularly if one is interested in determining whether they reflect differences in the sociocultural environment, such as nursing

habits in these countries or differences of a genetic-constitutional nature. Several types of studies come to mind, *e.g.*, a comparison of the frequency of breast cancer among Japanese-Americans who were born in Japan with that among American-born women of Japanese descent. Similar studies of those of Finnish birth and descent would be important.

TABLE 1.—*Average annual age-adjusted death rates (per 100,000 population) for breast cancer among females (1954–1955) and males (1950–1957) for selected foreign countries and the United States**

Country	Female	Male
Denmark	24.2	0.23
England and Wales	23.2	0.25
Canada	22.2	0.22
United States (white)	21.7	0.25
United States (nonwhite)	19.1	0.35
Sweden	18.9	0.15
Germany (Fed. Repub.)	15.1	0.21
France	14.0	0.34
Finland	11.7	0.04
Japan	3.9	0.07

*Adjusted with standard population used by Segi (1959) from the age distribution of the total population of 46 countries, 1950.

Stimulated by animal investigations indicating the influence of genetic factors, many familial studies of female breast cancer have been carried out. These studies have consisted of determining the prevalence of breast cancer among the relatives of a group of breast cancer patients and comparing it with an expected number, estimated from some type of control group. In table 2, we have summarized the results of these studies with respect to the relative prevalence or mortality of breast cancer among mothers and sisters of breast cancer patients. Of these 8 studies, 6 showed the presence of familial aggregation and 2 did not. In those studies showing familial aggregation, there is a fairly consistent finding of a twofold excess frequency of breast cancer among mothers and sisters. On the whole, these results suggest the influence of genetic factors. However, one must also keep in mind the possibility that familial aggregation may be a manifestation of environmental factors common to members of the same family.

TABLE 2.—*Summary of results of studies of familial aggregation of breast cancer*

Author, Year	Number of cases studied	Prevalence or mortality of breast cancer among:					
		Mothers of cases			Sisters of cases		
		Observed	Expected	Ratio: O/E	Observed	Expected	Ratio: O/E
Jacobsen, 1946	200	21	7	3.0	13	5	2.6
Penrose *et al.*, 1948	About 300	25	11	2.3	23	7	3.3
Passey *et al.*, 1952	585	23	20	1.1	----	----	----
Smithers, 1948	556	29	13.9	2.1	----	----	----
Woolf, 1955	200	4	2.1	1.9	8	3.2	2.5
Anderson *et al.*, 1958	544	9	7.7	1.2	28	12.4	2.3
Murphy and Abbey, 1959	200	7	3	2.3	2	6	0.3
Macklin, 1959	295	11	5.6	2.0	14	5	2.8

The degree of familial aggregation can be evaluated in the light of the other epidemiological data that we have reviewed. These data, viewed totally, are consistent with indicating that endogenous, hormonal factors are of etiologic importance and one would naturally expect these factors to be under genetic control. Therefore, it would seem to us that we should expect a greater degree of familial aggregation than actually has been observed, particularly since similar degrees of familial aggregation have been noted in studies of cancer sites where the epidemiological evidence suggests that environmental factors may play a considerable role, such as in the case of gastric cancer. This relative inconsistency of pattern awaits explanation.

In addition to the specific individual factors already discussed, there are several others that have been reported as being associated with the occurrence of breast cancer among females. It has been observed that the incidence and/or mortality rates from breast cancer are higher in the upper socioeconomic groups than in the lower ones (Graham *et al.*, 1960), higher in the urbanized areas of the United States (Levin *et al.*, 1960), higher among the Jewish group in New York City than among other religious groups (Newill, 1961), higher among the native than the foreign born in the United States (Haenszel, 1961), and higher among those with benign breast conditions (Lewison and Lyons, 1953). Further study of these differences is needed.

With regard to the suggestive etiologic importance of endocrine factors, it is of interest that Sommers (1955) and Brachetto-Brian and De Srulijes (1954) observed histologic changes of the ovary suggesting the existence of hyperestrinism in women with breast cancer. Their studies were based on autopsy material, limiting the inferences that can be made. There have been several reports that disordered thyroid function, as shown by a higher frequency of goiter, may be related to breast cancer (Bogardus and Finley, 1961; Liechty *et al.*, 1963). Again, many of these studies are subject to methodological objections, although further, better-controlled observations are warranted. It would appear fruitful to compare a profile of the endocrine status of breast cancer patients with that of controls. Attempts at such studies have been limited by methodological problems which may be resolved in the near future. As Bulbrook and Strong (1958) have indicated recently, whatever studies have been conducted have yielded results which in many respects are disappointing. Perhaps, the solution of the methodological problems will eventually permit answers to some of the questions that have been raised.

In considering possible epidemiological approaches to a disease problem which appears principally to be a result of endogenous factors, it occurred to us that it would be worthwhile to study the characteristics of male patients with breast cancer. Because of the rarity of this condition among males and because of the seeming importance of female-type hormones, these individuals might be strikingly

different. During 1961 and 1962, such a study was conducted by Schottenfeld and Diamond (1963); the details of the study will be reported elsewhere but we will briefly summarize the results here.

It was possible to interview 53 male breast cancer patients who had been treated during 1955–1962 in several clinics in many of the large cities in the U.S. For purposes of comparison, 53 patients with gynecomastia and 53 with cancer of the colon, matched by age and race were also interviewed. There was an excess frequency of Jews, single men, and those with later ages at first marriage among the breast cancer patients than among the others; although these were not statistically significant they were consistent with epidemiological data for females with breast cancer.

Of greater interest were the findings with respect to a history of orchitis, orchiectomy, the receipt of hormonal therapy, family history, X-ray exposure and antecedent benign breast disease. The results of the comparisons for these specific factors which may be of carcinogenic importance, are presented in table 3. In general, the number of patients with any of these characteristics is rather small, but there is an excess frequency of them among the breast cancer patients as compared to the control groups. Since several patients had one or more of these characteristics, the total number of patients who had one or more of these characteristics was computed; 22 percent of the breast cancer patients gave such a history as compared to 8 percent of the gynecomastia and 2 percent of the colon cancer group.

TABLE 3.—*Summary of history of characteristics among male patients with breast cancer, gynecomastia, and colon cancer*

Study group	Total patients	Number of patients with history of—						Total of patients with one or more characteristics	Percent with one or more characteristics
		Orchitis	Orchiectomy	Hormones (thyroid, cortisone, testosterone, stilbesterol)	Family history of cancer of male breast	Exposure to X-ray to chest	Antecedent benign breast disease		
Breast cancer	53	4	1	3	2	2	4	11	22
Gynecomastia	53	1	0	3	0	0	0	4	8
Colon cancer	53	0	0	1	0	0	0	1	2

Of specific interest from an endocrine viewpoint are the four patients with a history of orchitis and one with an orchiectomy. In view of the other epidemiological evidence the most acceptable interpretation is that orchitis produced a certain degree of testicular atrophy resulting in hormonal changes, which in turn played an etiological role in the development of the breast cancer. However, one cannot exclude the possibility that those with a history of orchitis may have had a concurrent infection of breast tissue which directly produced the neoplastic transformation.

The history of familial aggregation in two patients is extraordinarily high in view of the rarity of the disease. However, interpretation of this finding is difficult since it may reflect either genetic factors or environmental factors, such as the familial aggregation of infectious diseases producing orchitis. The exposure to chest X-rays is somewhat unexpected since this is the first instance, insofar as we are aware, of the suggestion that X-rays may be implicated in breast cancer. It should be noted that these represent exposures to therapeutic X-rays, probably of high doses.

Any generalization of these results should be made cautiously in view of the limited nature of these data. But they are sufficiently interesting and provocative to serve as a basis for serious consideration of further research along similar lines.

The review of some of the salient features of breast cancer illustrates the epidemiological approach to a disease in which endogenous factors appear to be of major etiological importance. Despite the complexities of the relationships, we discussed breast cancer first, since in a conference devoted to genetic contributions to epidemiologic studies, we wanted to indicate that epidemiology is not necessarily limited to a consideration of environmental determinants of disease, but attempts to achieve an integrated, holistic concept.

Now, we would like to consider some aspects of the epidemiology of a cancer site where the epidemiological evidence indicates that the environment plays a major etiological role. There is no need to review in detail the evidence that implicates cigarette smoking, exposure to certain selected occupations, and an urban factor which may be air pollution, as being of importance in the etiology of lung cancer. We will limit our discussion to a brief and selective review of some of the available data. Many retrospective studies have indicated that a larger proportion of lung cancer patients are cigarette smokers than a comparative group of controls, and of those who smoke, a larger proportion are heavy smokers. For various methodological reasons it was considered desirable to confirm this association by means of a prospective study. Such a study was initiated in 1952 by Hammond and Horn (1958) of the American Cancer Society. Histories of smoking habits were obtained on about 190,000 men aged 50 to 69 years; these men were traced for 44 months. Death certificates were obtained on all who died and additional medical information was obtained on those who were reported to have died of cancer. Altogether about 12,000 deaths were reported, of which 2,249 were attributed to cancer. The death rates for verified bronchogenic carcinoma by amount of cigarettes smoked are presented in figure 4. We note the increasing rate of dying from lung cancer with increasing amount of cigarettes smoked. Of additional interest is the observation that the death rate from lung cancer was lower for those who had been regular cigarette smokers in the past but were not smoking at the time of the study.

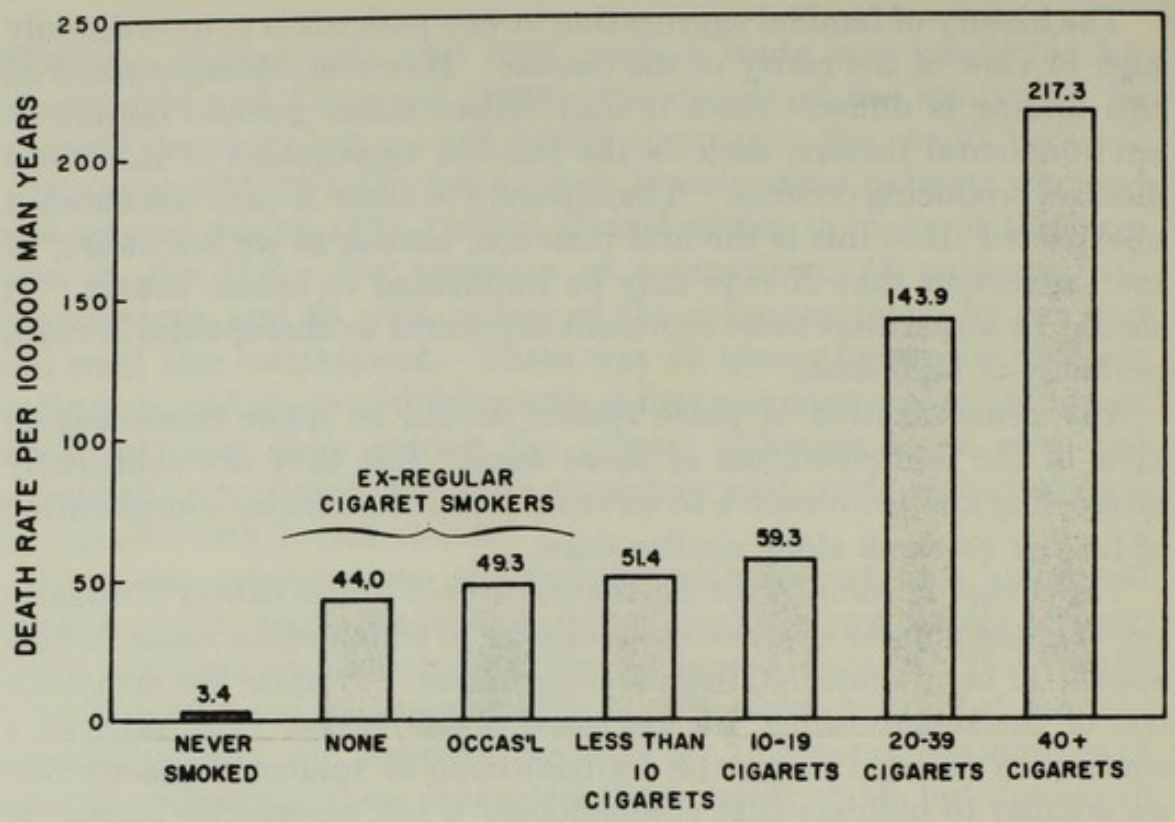
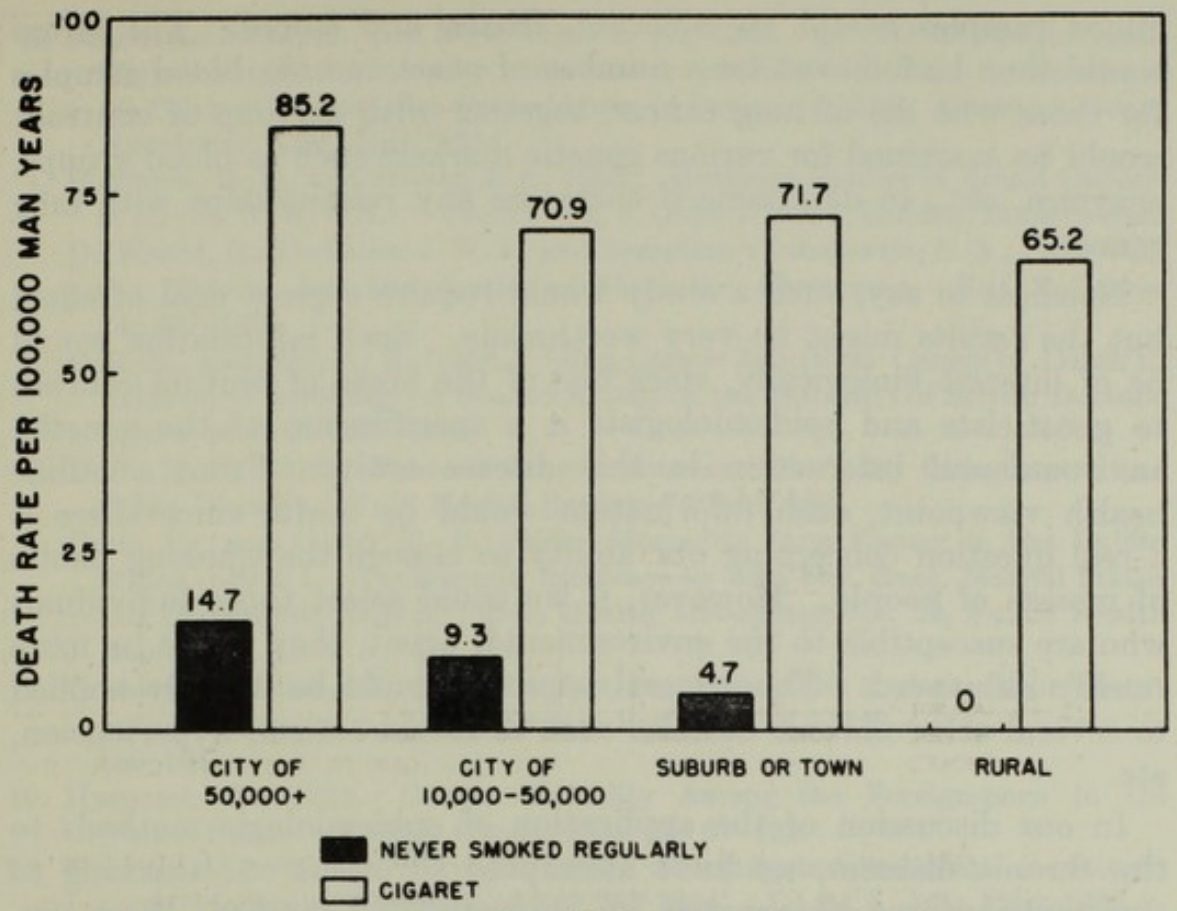


FIGURE 4.—Age-adjusted death rates of well-established cases of bronchogenic carcinoma (exclusive of adenocarcinoma) by number of cigarettes smoked daily.

Two additional prospective studies by Doll and Hill (1956) and by Dorn (1959) had similar results.

Hammond and Horn were also able to look at the influence of a possible urban factor. In figure 5, there are presented the death rates from lung cancer among cigarette smokers and nonsmokers in rural and urban areas of various sizes. The lung cancer death rates were a great deal higher among regular cigarette smokers than among nonregular smokers, regardless of whether they lived in urban or in rural areas. However, for both cigarette smokers and nonsmokers, the death rates were higher in the more urbanized areas than in the rural ones. But the data do also indicate that the contribution from smoking is greater than that from urbanization. In addition to these two environmental agents—smoking and urbanization—there is evidence that certain occupational exposures increase the lung cancer risk. But the contribution of this to the total amount of lung cancer is relatively small.

Recently, it has been shown (Tokuhata and Lilienfeld, 1963), as Tokuhata will report this afternoon, that there is a familial factor influencing the occurrence of lung cancer. This can be most reasonably interpreted as reflecting a genetic component which would influence an individual's susceptibility to the major environmental agent—cigarette smoking.



Age standardized death rates due to well-established cases of bronchogenic carcinoma (exclusive of adenocarcinoma) by urban-rural classification. Rates for cigarette smokers are compared with those for men who never smoked regularly.[43]

FIGURE 5.—Age-adjusted death rates of well-established cases of bronchogenic carcinoma (exclusive of adenocarcinoma) for regular cigarette smokers and nonsmokers by urban-rural classification.

To provide some idea of the relative risk associated with cigarette smoking, it has been estimated that a person who smokes two or more packs of cigarettes a day has a lifetime probability of dying from lung cancer of about 1 in 10 as compared to a nonsmoker whose lifetime probability is about 1 in 700. This suggests that there may be a segment in our population who are susceptible to this particular environmental agent. This is not an unreasonable assumption in view of what we know about the biology of disease. It is also reasonable to assume that this susceptibility reflects genetic factors, particularly in the light of the results of the familial studies.

From an epidemiological viewpoint and also from a genetic viewpoint it would be of interest to determine the specific factors or indexes of susceptibility. Unfortunately, there is no theoretical basis for designing a study to determine such factors. The only type of approach that could be adopted is a "shotgun" one in which a large group of males in appropriate age groups, perhaps 40,000 or 50,000 such individuals, would be enrolled in a followup study. Initially, information would be obtained with respect to various demographic characteristics, smoking habits and other environmental exposures.

Blood samples would be collected, frozen and stored. The group would then be followed for a number of years and the blood samples for those who die of lung cancer, together with a group of controls, would be examined for various genetic markers such as blood groups, enzymes, etc., to determine if there are any relationships with lung cancer.

Needless to say, such a study would require a great deal of effort but the results might be very worthwhile. Such information would be of interest biologically, since one of the areas of mutual concern to geneticists and epidemiologists is a specification of the genetic-environmental interaction in this disease entity. From a public health viewpoint, such information would be useful since there is a real question concerning our ability to change the smoking habits of masses of people. However, if we could select those individuals who are susceptible to the environmental agent, they might be more readily influenced. This general approach could be usefully applied to several other chronic diseases such as breast cancer, hypertension, etc.

In our discussion of the application of epidemiologic methods to the chronic diseases, we have attempted to orient our thinking to the contributions of genetics to epidemiologic studies. However, we should not lose sight of the other side of the coin—the contributions that epidemiologic methods could make to genetic studies. Time does not permit an expansion of this theme. Briefly, we would like to call attention to the need by geneticists for reliable data on the incidence of disease as a basis for studies of linkage and modes of inheritance. There is also a need for determining the environmental causes of chromosomal aberrations, such as those producing nondisjunction, and of mutations. No doubt there will be some discussion of these issues in the workshops this afternoon.

In this rather brief exposition of the use of epidemiologic concepts in the chronic diseases, we have tried to emphasize that the epidemiologist, who is interested in studying the chronic diseases, depends upon the accumulation of knowledge of various kinds, clinical, genetic, epidemiological and others. He is interested in the totality of the data obtained by different approaches, from which a pattern emerges and serves as a basis for deriving fairly valid inferences concerning the biology of the disease in human populations.

REFERENCES

1. Anderson, V. E., Goodman, H. O., and Reed, S. C., 1958. Variables Related to Human Breast Cancer. Minneapolis, University of Minnesota Press.
2. Bogardus, G. M., and Finley, J. W., 1961. Breast Cancer and Thyroid Disease. Surgery 49: 461–468.

3. Brachetto-Brian, D., and De Srulijes, L. K., 1954. Le role de l'ovaire dans l'hyperestrogenisme chez les femmes avec cancer la mamelle. International (6) Cancer Congress, Sao Paulo (Abstr. of Papers): p. 249. Quoted in CA 4: Newsletter XIV, Nov. 1954.
4. Bulbrook, R. D., and Strong, J. A., 1958. Hormone Studies in Breast Cancer, in Raven, R. W. (Ed.). Cancer, Vol. 6, Chpt. 7: 215. London, Butterworth.
5. De Waard, R., De Laive, J. W. J., and Baanders-Vanhalewijn, E. A., 1960. On the Bimodal Age Distribution of Mammary Carcinoma. Brit. J. Cancer 14: 437–448.
6. Doll, R., and Hill, A. B., 1956. Lung Cancer and Other Causes of Death in Relation to Smoking—A Second Report on the Mortality of British Doctors. British Med. J. 1071–1081.
7. Dorn, H. F., 1959. Tobacco Consumption and Mortality from Cancer and Other Diseases. Pub. Health Reports 74: 581–593.
8. Dorn, H., and Cutler, S. J., 1955. Morbidity from Cancer in the United States. Part I. Variation in Incidence by Age, Sex, Race, Marital Status and Geographic Region, Public Health Monograph No. 29, Public Health Reports.
9. Graham, S., Levin, M., and Lilienfeld, A., 1960. The Socioeconomic Distribution of Cancer in Various Sites in Buffalo, N.Y., 1948–1952. Cancer 13: 180–191.
10. Haenszel, W., 1961. Cancer Mortality Among the Foreign-born in the United States. J. Nat. Cancer Inst. 26: 37–132.
11. Hammond, E. C., and Horn, D., 1958. Smoking and Death Rates—Report on 44 Months of Followup of 187,783 Men. J.A.M.A. 166: 1294–1308.
12. Jacobsen, O., 1946. Heredity in Breast Cancer: A Genetic and Clinical Study of Two Hundred Probands. London, H. K. Lewis.
13. Kamoi, M., 1960. Statistical Study on Relation between Breast Cancer and Lactation Period. Third Report. Study through Relative Liability. Tohoku J. Exper. Med. 72: 72–77.
14. Levin, M. L., Haenszel, W., Carroll, B. E., Gerhardt, P., Handy, V. H., and Ingraham, S. C., II, 1960. Cancer Incidence in Urban and Rural Areas of New York State. J. Nat. Cancer Inst. 24: 1243–1257.
15. Lewison, E. F., and Lyons, J. G., Jr., 1953. Relationship between Benign Breast Disease and Cancer. A.M.A. Arch. Surg. 66: 94–114.
16. Liechty, R. D., Hodges, R. E., and Burket, J., 1963. Cancer and Thyroid Functions. J.A.M.A. 183: 30–32.
17. Lilienfeld, A. M., 1956. The Relationship of Cancer of the Female Breast to Artificial Menopause and Marital Status. Cancer 9: 927–934.
18. Lilienfeld, A. M., 1959. Epidemiologic Methods and Inferences, in Hilleboe, H. E., and Lammiore, G. W., editors: Preventive Medicine, Philadelphia, W. B. Saunders Company.
19. Lilienfeld, A. M., 1961. Some Epidemiologic Aspects of Cancer of the Breast. Proceedings Fourth National Cancer Conference, Philadelphia, J. B. Lippincott Co., pp. 215–223.
20. Lilienfeld, A. M., 1963. The Epidemiology of Breast Cancer. Presented at Annual Meeting, Amer. Assn. Cancer Research, May 23, 1963.
21. MacMahon, B., and Feinleib, M., 1960. Breast Cancer in Relation to Nursing and Menopausal History. J. Nat. Cancer Inst. 24: 733–753.
22. MacMahon, B., Pugh, T. F., and Ipsen, J., 1960. Epidemiologic Methods. Boston, Little, Brown and Co.
23. Macklin, M. T., 1959. Comparison of the Number of Breast Cancer Deaths Observed in Relatives of Breast Cancer Patients and the Number Expected on the Basis of Mortality Rates. J. Nat. Cancer Inst. 22: 927–951.

24. Murphy, D. P., and Abbey, H., 1959. Cancer in Families. (A Study of the Relatives of 200 Breast Cancer Probands.) Cambridge, Harvard Univ. Press.
25. Newill, V. A., 1961. Distribution of Cancer Mortality among Ethnic Subgroups of the White Population of New York City, 1953–58. J. Nat. Cancer Inst. 26: 405–417.
26. Passey, R. D., Waiman, M., Armstrong, E. C., and Rhodes, J., 1952. Heredity in Human Breast Cancer. Acta Unio Internat. Contra Cancrum 8: 184–186.
27. Penrose, L. S., Mackenzie, H. J., and Karn, M. N., 1948. A Genetical Study of Human Mammary Cancer. Eug. 14: 234–266.
28. Schottenfeld, D., Lilienfeld, A., and Diamond, H., 1963. Some Observations on the Epidemiology of Breast Cancer among Males. Am. J. Public Health (in press).
29. Segi, M., Fuzisaku, S., Kurihara, M., Narai, Y., Ito, M., and Ogata, Y., 1959. Age-adjusted Death Rates from Malignant Neoplasms for Selected Sites, by Sex, in 24 Countries, in 1952–53, 1954–55, and 1956–57. Sendi, Japan, Tohoku Univ. School of Medicine, Dept. of Public Health.
30. Smithers, D. W., 1948. Family Histories of 459 Patients with Cancer of the Breast. Brit. J. Cancer 2: 163–167.
31. Sommers, S. C., 1955. Endocrine Abnormalities in Women with Breast Cancer. Lab. Invest. 4: 160–174.
32. Stocks, P., 1955. Social Status in Relation to Carcinoma of the Breast. Schweiz Z. Path. Bakt. 18: 706–717.
33. Tokuhata, G., and Lilienfeld, A., 1963. Familial Aggregation of Lung Cancer in Humans. J. Nat. Cancer Inst. 30: 289–311.
34. Versluys, J. J., 1955. Marriage, Fertility, and Cancer Mortality of the Specifically Female Organs: Mammary Carcinoma. Brit. J. Cancer 9: 239–245.
35. Woolf, C. M., 1955. Investigations on Genetic Aspects of Carcinoma of the Stomach and Breast, in Univ. of California Publications in Public Health. Vol. 2, p. 265. Berkeley, Univ. of California Press.
36. Wynder, E. L., Bross, I. J., and Hirayama, T. A., 1960. A Study of the Epidemiology of Cancer of the Breast. Cancer 13: 559–601.
37. Yerushalmy, J., and Palmer, C. E., 1959. On the Methodology of Investigations of Etiologic Factors in Chronic Diseases. J. Chron. Dis. 10: 27–40.

PART II

The Evaluation of Genetic Factors in Selected, Illustrative Diseases

Diabetes Mellitus[1]

By JAMES V. NEEL,[2] M.D., Ph. D., STEFAN S. FAJANS,[3] M.D., JEROME W. CONN,[3] M.D., and RUTH T. DAVIDSON,[2] B.S., *University of Michigan Medical School, Ann Arbor, Mich.*

INTRODUCTION

Diabetes mellitus is in many respects a geneticist's nightmare. As a disease, it presents almost every impediment to a proper genetic study which can be recognized.

Firstly, the nature of the basic defect is unknown, and as long as this is true, the possibility that the disease may be heterogeneous hovers overhead. Thus, in recent years it has been variously postulated that the primary defect in the majority of cases is an abnormal insulin molecule (Conn and Fajans, 1961), an excess of insulin antagonists (Vallance-Owen and Lilley, 1961), an initial overproduction of insulin followed by excessive antagonist activity (Neel, 1962), or an intrinsically higher rate of release of fatty acids and ketone bodies for oxidation, with, as one consequence, an impaired sensitivity to insulin (Randle, Garland, Hales, and Newsholme, 1963). The further possibility has been considered that the primary defect has nothing to do with insulin at all, but results from an abnormality in the basement membrane of the small blood vessels (review in Siperstein, in press).

Secondly, the frequency of this disorder in the general population is not well defined, a problem if one wishes to test hypotheses which draw on the concept of gene frequency. Thus, from household interviews of the civilian, noninstitutionalized population, conducted in 1,900 sampling units throughout the United States, the frequency of

[1] This investigation was supported in part by Grants A-888 and AM 04108-03 GEN from the National Institutes of Health. The participation in this study of Dr. Lee Schacht, Dr. T. Edward Reed, and Dr. Richard Post is gratefully acknowledged.

[2] Department of Human Genetics.

[3] Department of Internal Medicine.

recognized diabetes mellitus was estimated (rates per 100 persons) to be as follows (U.S. National Health Survey, 1960):

Age	*Males*	*Females*
0–24	0. 11	0. 07
25–44	0. 49	0. 38
45–54	1. 12	1. 37
55–64	2. 52	3. 15
65–74	3. 44	5. 03
75+	3. 15	3. 88

On the other hand, when screening-type glucose tolerance tests, followed by more extensive studies when indicated, have been performed, significantly higher figures have been obtained. For example, Chesrow and Bleger (1955), after exclusion of known diabetics, found 64 "definite" and 24 "potential" diabetics among 1,000 residents of an infirmary who were over 60 years of age. Allowance for known and/or deceased diabetics would be expected to bring the total cumulative rate to close to 10 percent. Brandt and Tupper (1960) have reported even higher figures for inmates of county hospitals. Obviously, these figures while useful, cannot be used for normative values. Nonspecific factors which may affect glucose tolerance, such as decreased mobility, increased bed rest, chronic illness, decreased food intake, and other factors, frequently operative in older people, may contribute to the high incidence of abnormal carbohydrate metabolism in these series.

More to our needs, in a large scale survey in Georgia, 6.7 percent of persons aged 60–69 and 7.4 percent of persons over 70 were found to have significantly lowered glucose tolerance. Both Negro and Caucasian subjects were included; highest rates were encountered in Negro females and lowest in Negro males. These figures include known diabetics (Petrie, McLoughlin, and Hodgins, 1954). In a health appraisal program at the University of Michigan, among the first 548 subjects evaluated, 90 percent male, with an average age of 56.5 years, there were 10 or 1.8 percent with known diabetes but an additional 59 or 10.8 percent were thought to have previously unsuspected diabetes (Rush and Tupper, 1960). Criteria for diagnosis in 42 of this latter group were a blood sugar above 160 mg percent at 1 hour, and above 120 mg percent at 2 hours. In 17 of the 59 subjects, a diagnosis of "diabetes" was made solely on the basis of the 2-hour postprandial (100 gm glucose) blood sugar test. These examinations were made without the use of a preparatory diet and on the same day that other routine tests such as chest and gastrointestinal roentgenograms were carried out. In a comparison of the 2-hour postprandial blood sugar determined in this fashion with the standard oral glucose tolerance tests performed by the same authors in 110 patients, it was noted that the 2-hour value of the glucose tolerance tests was lower (mean 31 mg percent, range 1–117 mg percent) in approximately 75

percent of this group. Thus, the figure of 10.8 percent for "persons with unsuspected diabetes" appears to be high. It should also be noted that the 10.8 percent figure was "derived from the examining physician's final diagnosis of each patient concerned. No attempt was made in any of these cases to impose an absolutely standardized interpretation of the laboratory results."

Whatever the true incidence of diabetes may be, these findings suggest that a rather considerable number of persons with what we must at the very least term "abnormal carbohydrate tolerance" and at the most mild diabetes, will ordinarily live out their days without this fact becoming evident. The results of twin studies provide further evidence to this effect.

Thus, Then Bergh (1938) studied 35 monozygous and 35 dizygous twin pairs age 43 or over, ascertained through the fact of diabetes in one member of the pair. Fifteen monozygous and two dizygous pairs were concordant. He was able to administer glucose tolerance tests to 11 of the apparently normal monozygous co-twins—all the tests were felt to be abnormal. A followup study would be of great value, to see for how many pairs concordance in the clinical sense ultimately developed. Be that as it may, it is apparent that the evolution of diabetes mellitus follows a wide variety of courses, with, even in the advanced years, many persons with the predisposition not having overt disease.

Thirdly, and closely related to the foregoing, is the fact that it is by no means clear whether diabetes mellitus should be regarded as a distinct, qualitative departure from normal glucose metabolism, or whether, rather, the clinical entity is the "tail" of a continuous distribution, with some individuals in this tail exhibiting secondary complications which contribute to the impression of a discrete disease. There has never been the extensive, detailed population survey which would permit an opinion on this point, although we shall later present some pertinent data. To date the geneticist, taking his clue from the clinician, has usually approached diabetes as if it were a discrete entity. A rather different approach is indicated if there is a continuous, unimodal distribution to the basic defect(s).

Fourthly, now, it is clear that the frequency with which the disease is diagnosed appears to be strongly influenced by environmental variables, especially the general nutritional level. For instance, the frequency of diabetes appears to differ among repatriate Jewish groups in Israel (Cohen, 1963). The ultimate strength accorded this finding as evidence for the action of environmental factors will be strongly influenced by information still unfolding regarding intermarriage in the countries of dispersion (discussion in Goldschmidt, 1963). A further uncertainty pertains to death rates from diabetes prior to repatriation.

Fifthly, the disease may become apparent at any age from 1 to 100. The geneticist at any given time is thus working with only a fraction of the possible diabetics, and must accordingly, if he wishes to think in terms of ratios, make adjustments for this fact, adjustments the more difficult because the published age-of-onset and prevalence curves are at variance with one another and the sexes are not equally affected by the disease (Spiegelman and Marks, 1946; Joslin *et al.*, 1959; U.S. National Health Survey, 1960).

Sixthly, until recently the so-called juvenile diabetics seldom reproduced. Attempts to determine whether the genetic basis for juvenile diabetes might differ from that of the adult type thus have been handicapped by a scarcity of certain mating types rather critical to any tests of hypotheses.

Seventhly, and finally, there are well-known problems in the reproducibility of certain diagnostic procedures. Thus, when the position of individual points on the GTT is subject to statistical analysis, considerable variability is observed in the results of repeated tests. For instance, Hayner, Waterhouse, and Gordon (1963), in a series of observations on 28 prisoners on standard prison fare, with few exceptions in the fifth decade, have found the following absolute differences and standard deviations for corresponding values on glucose tolerance tests repeated one week apart: 0 hours—4.4 ± 3.8 mg percent; ½ hour—16.8 ± 15.2 mg percent; 1 hour—20.0 ± 18.0 mg percent; 1½ hours—14.1 ± 12.5 mg percent; 2 hours—11.9 ± 13.5 mg percent; and 3 hours—16.1 ± 14.7 mg percent. In spite of the magnitude of these differences and standard deviations between replicate observations at the various points of the two curves, it needs, for clinical and diagnostic purposes, to be emphasized that both members of the pair of glucose tolerance tests either fell well within the limits defined as normal (see below), or abnormal (1 case) for 27 of the 28 prisoners. The individual with the mildly abnormal glucose tolerance tests had a family history of diabetes mellitus. One individual in this series had one normal test and one showing an elevation of the 1-hour level of the glucose tolerance test.

Why, you may wonder after this dreary recital, would any geneticist ever venture into this obviously genetically unprofitable arena? The answer is very simple. This is a relatively common and potentially fatal disorder in which there is much evidence for a familial predisposition. Not only is it a tempting target for the investigator, but its frequency raises some fascinating problems not encountered in the study of such rare metabolic disorders as galactosemia or albinism. Assuming a genetic basis, how did a disease with such an impact on reproduction reach such relatively high frequencies? Why hasn't natural selection reduced the frequency of the diabetic genotype(s) to that of most other serious diseases in whose etiology genetic factors seem of importance? There was a further, practical reason

for our venturing into this field some ten years ago, namely, the development of the cortisone-conditioned glucose tolerance test (Fajans and Conn, 1954), with the possibility of a very useful handle for the identification of a considerable proportion of individuals with the diabetic genotype long before the onset of clinical disease.

REVIEW OF LITERATURE

That diabetes mellitus occurs more frequently in some families than might be expected on the basis of chance alone has been recognized for centuries, although Allan (1933), Pincus and White (1933; 1934a, b), and Levit and Pessikova (1934) were apparently the first to place this impression on a sound statistical basis. Since these familial concentrations were observed within neighboring socioeconomic levels, in the absence of any identifiable environmental factors, a genetic etiology seemed likely. The many studies of the past three decades designed to elucidate the most likely genetic basis for this nonrandom distribution have in recent years been reviewed by Grunnet (1957), Steinberg (1959, 1961), Lamy, Frézal, and Rey (1961), and Simpson (1962). At one time or another, such diverse simple genetic mechanisms have been suggested as irregular dominance, recessiveness, and a heterozygous-homozygous relationship, wherein the heterozygote tends to develop the mild diabetes of the later years and the homozygote the more severe diabetes of the earlier years. Each of these theories continues to have its proponents, but recently a number of investigators, impressed by the difficulty of explaining the facts with a single pair of genes, have begun to consider a more complex, multigenic type of inheritance. The fact that able investigators could reach such diverse and/or uncertain conclusions reflects the difficulties in the study of the disease which were enumerated earlier.

The findings thought to suggest recessive inheritance fail to indicate whether the presumed recessive genes occur at only one or at several different loci. Now, if all cases of a disease are due to recessive inheritance involving the same locus, all the offspring of two affected parents should themselves be affected. No other reasonably simple genetic hypothesis leads to this expectation. Accordingly, a rather critical test of at least one of the hypotheses regarding diabetes would appear to be afforded by the findings in the offspring of conjugal diabetics. If only one locus is concerned, then all the offspring should eventually develop diabetes. If, however, more than one locus is concerned, or if heterozygotes for this "recessive" gene sometimes develop diabetes, then the offspring would not all be expected to develop diabetes.

Post (1962) has recently presented an analysis of such critical material, based in part on the data of West (1960), and in part on pedigrees collected in connection with the study to be described shortly. He reported that in a total material of 161 offspring ranging in age from the first to the seventh decade, 67 from West's data, and 94 from our own study files, there were 13 diabetics. Expectation for a group with this age distribution, if all were destined ultimately to develop diabetes, was 8.7. The difference between observation and expectation was not statistically significant.

In this observation, it is of the utmost importance that ascertainment be solely through the parents or that, if otherwise, less adequate allowance be made for this fact. In reviewing both the paper of West (1960) and the data utilized by Post (1962), in preparation for this discussion, we have been impressed that problems of ascertainment are still with us. Thus, in the work by West (1960), "all the subjects tested with a family history of diabetes came to our attention through their diabetic parents or their nondiabetic siblings." The frequency of ascertainment of the latter type, through nondiabetic siblings, is not stated. To the extent that it occurred, it would tend to produce a bias in the number expected, in the direction of a relative deficiency of affected.

Reexamination of the kindreds on which Post (1962) based his report also reveals some probable bias, in that in four sibships it seems very likely that ascertainment was significantly influenced by the fact that the conjugal diabetics had produced an affected offspring, i.e., that ascertainment was at least in part through the offspring. The simplest way to correct for this fact is to omit the offspring through whom ascertainment occurred (diabetic or normal) from consideration. Furthermore, one of the sibships was removed for diagnostic reasons. After these corrections, and with the addition of two new sibships, our own material contains 99 offspring of conjugal diabetics, of whom 6 were already diabetic. Combined with West's material, with the reservations stated above, the total now becomes 166 offspring, 11 affected. Expectation in a group of this age structure on the hypothesis that eventually all will be affected is 6.6; $X^2_{(1)}$ is 1.9 and the difference is not significant. I should emphasize that we have scored as diabetic in this accounting only those whose diabetes was known when the family came to attention, because this gives us diagnostic comparability with the figures of the surveys on which U.S. frequencies are based.

The numbers involved are still quite small. Although the data are consistent with the hypothesis that all the offspring will in time be affected, they do not exclude the hypothesis that some of the children will remain normal. Thus, let us assume for the sake of the argument that these affected parents are actually all heterozygous for a gene

which can produce diabetes in either homozygote or heterozygote.* Then three-quarters of their children would be expected to develop diabetes but one-quarter should escape the disease. This would reduce expectation in this sample to 4.98 individuals. The difference between observation and expectation is still not significant ($X^2_{(1)}=3.5$, $.10>P>.05$); we cannot exclude such a different hypothesis on the basis of the available data!

Earlier we considered the possibility that a substantial proportion of those with the genetically determined predisposition to diabetes will fail to develop the disease during their lifetimes. If this is so, then even among the offspring of conjugal diabetics we might expect a considerable proportion of the children to remain clinically unaffected. This makes the earlier mentioned excess of affected over expectation among the offspring of conjugal diabetics, even though nonsignificant in these limited numbers, potentially somewhat puzzling. In this connection, we should recognize the possibility that our hospital sample is weighted for severity of diabetes, with a possible effect on the figures.

PLAN OF PRESENT STUDY

Some 10 years ago we initiated a study on the genetics of diabetes which was designed to take advantage of the possibility that the individual predisposed to diabetes might be recognizable years before the onset of clinical disease by his response to a standardized dose of cortisone. Standard glucose tolerance curves (GTT), the great majority repeated after the administration of cortisone, have now been studied in some 103 controls and 573 apparently healthy, first degree relatives of diabetic patients seen at the University of Michigan Hospital. The controls are for the most part medical students, physicians, technicians, and nurses. Blood sugars were determined by the Nelson modification of the Somogyi method. In this presentation we propose first to utilize a portion of these data to test the hypothesis of simple recessive inheritance for diabetes mellitus, a hypothesis recently found by Steinberg (1959, 1961) to be "the only one consistent with the several large samples of family data which have been published." Since the data do not fully conform to expectation on this hypothesis, we will then explore alternative explanations of the findings.

We are at the outset badly in need of some definitions and standards. Diabetes mellitus, as previously discussed (Jackson and Woolf, 1957; Fajans and Conn, 1959) and as we shall shortly demonstrate again,

*The distribution of homozygotes and heterozygotes in populations being what it is, it is of course very unlikely all these parents would be heterozygotes; the assumption is made solely for illustrative purposes.

is not a disease of sudden onset, but one in which there have usually been distinct and progressive metabolic abnormalities for many years before the onset of overt disease. There is a wide range of response in the "normal" glucose tolerance curve, and an equally wide range of variation in the impairment of glucose metabolism in the so-called diabetic. What kinds of definitions and boundaries can we set up for purposes of identification?

In this presentation, in addition to "normal" people, we propose to recognize four classes of people who may fit into stages in the natural history of diabetes. Frank, florid, overt, or symptomatic diabetes is the most advanced of these stages. This stage is represented by the individual whose impairment of glucose metabolism or "complications" of the disease have brought him to medical attention. The next term to be used is latent, or asymptomatic but clinically detectable, diabetes. A latent diabetic is an individual who has no symptoms referable to the disease but in whom a diagnosis of diabetes can be made either by the finding of an elevated fasting blood sugar level or with the use of the standard glucose tolerance test.

To define latent diabetes, we must be able to define the normal. In table 1 are summarized by age-decade and sex the results of the glucose tolerance tests performed in the course of this study on 113 individuals between the ages of 10 and 39 *whose family histories were negative for diabetes mellitus*. It should be emphasized that these were not the usual cursory medical family histories but that detailed pedigrees were obtained on each individual! All subjects were on a 3-day "preparatory" diet containing 300 gm of carbohydrate per diem and were given a test meal consisting of 1.75 gm of glucose per kilogram of "ideal" body weight, calculated from the subject's height and age. In the calculation of ideal body weight all ages beyond 25 were equated to 25, since according to life insurance tables "ideal" body weight stays constant for a given height after this age. For the two decades for which significant numbers are available, it will be noted that male values average slightly higher than female, although for the 20–29 age interval (for which the most data are available) only one of the differences, at 2 hours, is significant at the 5 percent level. Incidentally, in this and other similar comparisons to follow it must be borne in mind that inasmuch as any individual glucose level at one time point is not independent of the level at the next or preceding time point, these t tests are not independent of one another. There is no clear age trend in this age interval. Since the sex difference is small in relation to the standard deviation, we have felt justified in deriving one composite curve for age 10–29 (average age 24.7 years), sexes combined, to be used in certain comparisons to follow; this curve is given as the last entry in the table. It is to be noted that the curves for ages 30 to 39 are the same or lower. Fajans and Conn (1954) and Conn and Fajans

(1961) have previously defined (in part on an experience with this same material) a diabetic glucose tolerance curve as one exhibiting a one-hour value of 160 mgm/100 ml or more, a 1½-hour value of 140 mgm/100 ml or more, and a 2-hour value of 120 mgm/100 ml or more. It will be seen that these values fall two standard deviations above the control curve, so that the definition appears suitably conservative *for this age group*. By contrast, a normal glucose tolerance curve was defined as one showing a one-hour value of less than 160 mgm/100 ml, a 1½-hour value of less than 135 mgm/100 ml, and a 2-hour value of less than 110 mgm/100 ml. In this connection, it must be emphasized that some of our controls may of course later develop diabetes despite the negative family histories; the control curve is thus an approximation. Be this as it may, we define a "latent" diabetic as an individual whose glucose tolerance curve is abnormal as defined above, and a "normal" individual as one whose curve falls within the defined limits. It is clear that the dividing line between "latent" and "overt" diabetes is arbitrary.

TABLE 1.—*Mean glucose tolerance tests by decade and sex, in individuals aged 10–39 with a negative family history for diabetes. The significance of all differences shown has been examined by a t-test, with one, two, and three asterisks indicating significance at the .05, .01, and .001 levels respectively. Variations in the numbers for the various test points in this and subsequent tables result from the fact that complete curves were not obtained on all individuals*

Age	Sex	Value	Hours after test meal						
			0	½	1	1½	2	2½	3
10–19	M	$\bar{x}$	88	111	70		71		76
		N	1	1	1		1		1
20–29	M	$\bar{x}\pm\sigma_{\bar{x}}$	82.3±0.9	130.5±3.1	106.4±3.7	91.7±3.0	90.1±1.9	87.2±1.9	85.2±2.2
		σ	6.2	21.6	26.9	20.3	13.7	13.7	16.0
		N	53	49	53	47	53	50	52
30–39	M	$\bar{x}\pm\sigma_{\bar{x}}$	84.5±1.8	127.1±4.6	104.3±7.1	84.7±5.1	88.5±4.9	90.4±3.1	79.9±4.7
		σ	7.0	17.9	27.6	19.8	19.1	11.8	18.1
		N	15	15	15	15	15	13	15
10–19	F	$\bar{x}$	79	116.3	100	88.3	101	85	77.7
		N	3	3	3	3	3	3	3
20–29	F	$\bar{x}\pm\sigma_{\bar{x}}$	80.3±0.9	125.8±3.8	98.9±3.9	88.1±3.1	84.3±2.2	83.3±2.6	80.4±2.8
		σ	5.3	22.2	23.3	18.4	12.8	15.0	16.4
		N	35	35	35	35	35	34	34
30–39	F	$\bar{x}\pm\sigma_{\bar{x}}$	79.7±2.2	126.2±7.0	93.7±6.9	76.8±8.8	78.1±5.0	83.0±8.4	83.7±6.5
		σ	5.3	16.7	16.8	19.6	12.1	18.8	16.0
		N	6	5	6	5	6	5	6
20–29	M–F	Diff.	2.0	4.7	7.5	3.6	*5.8	3.9	4.8
10–29	M,F†	$\bar{x}\pm\sigma_{\bar{x}}$	81.5±0.6	127.9±2.4	103.0±2.7	90.1±2.1	88.0±1.4	85.6±1.5	83.1±1.7
		σ	5.9	22.8	25.6	19.5	13.7	14.3	15.9
		N	92	88	92	85	92	87	90

†Mean age—24.7 years.

The next term to be defined is "abnormal cortisone-GTT," which *may* indicate the presence of "subclinical diabetes." In table 2 are summarized by decade and sex the findings of the cortisone-glucose tolerance tests of 101 nondiabetic individuals between the ages of 10 and 39 with negative family histories for diabetes. The test was administered as follows: after completion of the baseline glucose tolerance tests the subjects continued on the standard prep diet for lunch and dinner of that day. A second glucose test was administered

the following morning. At 8½ and again at 2 hours before that test was scheduled, the subject received cortisone acetate orally. If he weighed less than 160 lbs the dose was 50 mgm both times; if he weighed 160 lbs or more, the dose was 62.5. It should be emphasized that the dose of cortisone is arbitrary; a variety of other dose-time schedules could be employed but of course each would require a different definition of an "abnormal" response. It must also be stressed that the cortisone-glucose tolerance test should still be regarded as an investigative procedure and is not presently recommended for routine clinical use.

TABLE 2.—*Mean cortisone-glucose tolerance tests by decade and sex, in individuals aged 10–39 with a negative family history for diabetes. The significance of all differences shown has been examined by a t-test, with one, two, and three asterisks indicating significance at the .05, .01, and .001 levels respectively. Variations in the numbers for the various test points in this and subsequent tables result from the fact that complete curves were not obtained on all individuals*

Age	Sex	Value	Hours after test meal						
			0	½	1	1½	2	2½	3
10–19	M	$\bar{x}$	87	150	120		79		68
		N	1	1	1		1		1
20–29	M	$\bar{x}\pm\sigma_{\bar{x}}$	91.0±0.8	152.0±2.5	139.0±3.6	115.0±3.1	107.2±2.4	99.1±1.9	92.9±2.7
		σ	5.7	17.4	25.8	21.1	16.9	13.5	19.0
		N	50	45	51	46	51	50	50
30–39	M	$\bar{x}\pm\sigma_{\bar{x}}$	94.4±2.9	151.7±5.7	137.7±11.4	114.1±30.0	105.1±4.0	96.5±4.4	90.1±4.7
		σ	11.2	21.9	44.3	30.0	15.5	16.9	18.1
		N	15	15	15	15	15	15	15
10–19	F	$\bar{x}$	84.3	139.7	110.0	104.7	84.7	80.7	87.3
		N	3	3	3	3	3	3	3
20–29	F	$\bar{x}\pm\sigma_{\bar{x}}$	92.6±1.3	154.2±4.9	144.0±7.3	119.2±6.1	107.9±3.8	99.7±3.0	94.9±2.9
		σ	7.3	26.2	38.8	33.1	20.3	15.8	15.4
		N	29	29	29	29	29	28	29
30–39	F	$\bar{x}\pm\sigma_{\bar{x}}$	93.0	173.0	152.5	136.0	132.5	102.5	89.0
		N	2	2	2	2	2	2	2
20–29	M–F	Diff.	−1.6	−2.2	−5.0	−4.2	−0.7	−0.6	−2.0
10–29	M, F	$\bar{x}\pm\sigma_{\bar{x}}$	91.3±0.7	152.3±2.5	139.5±3.4	116.2±3.0	106.3±2.0	99.5±1.6	93.1±2.0
		σ	6.5	21.6	31.3	26.0	18.5	14.4	18.0
		N	83	78	84	78	84	81	83

As before, based on the most numerous age group (20–29) there is a suggestion of a sex difference, of small magnitude, and not significant for any of the "point" comparisons; but now it is the females who exhibit the higher values. There are only two age-sex groups in which the numbers are adequate to search for age trends—none are apparent. Accordingly, and again as a basis for comparisons to be described shortly, we have felt justified in deriving a single composite curve for both sexes, ages 10–29; this is the last entry in table 2. The combined curve for ages 30–39 is approximately the same as that given for ages 10–29. Comparison of this entry with the corresponding entry of table 1 illustrates the well-known effect of cortisone on glucose tolerance; variability, as measured by the standard deviation, is not significantly increased, but yet six of the seven standard deviations are larger after cortisone. Fajans and Conn (1954, 1961) have defined an abnormal response to cortisone as a 2-hour value in

excess of 140 mgm/100 ml; table 2 indicates that this is a properly conservative standard.

The last term to be defined is prediabetes. The prediabetic is an individual who has a (genetic) predisposition to the disease but in whom no diagnostic test in use today discloses an abnormality. Thus by this definition a prediabetic has a normal glucose tolerance test as well as a normal cortisone-glucose tolerance test.

As is apparent from the magnitude of the standard deviations, there is great individual variation in the glucose tolerance curve. How much of this is "random," of the type discussed in the opening section, and how much real, is not clear. Now, it is a matter of some interest, both for the results obtained in the controls and in the offspring of diabetics to be considered shortly, to determine whether there is a correlation between corresponding points on the standard and cortisone glucose tolerance curves of any given individual. This question is investigated in table 3. The correlations, of the order of .35–.40, are generally significant, but, viewed otherwise, under the conditions of these observations, only approximately 15 percent of the variation in the position of the cortisone tolerance curve at any one time would be removed by knowledge of the corresponding position on the standard test. While it seems quite likely that the true correlation is somewhat higher, being obscured by random noise in the system, the fact is that no investigator has published the type of observations necessary to validate that statement, namely, repetition within a short interval of both types of curves 3–5 times in a series of some 20 individuals.

TABLE 3.—*Correlations between corresponding points on the glucose tolerance curve and the cortisone-glucose tolerance curve, in individuals of age-sex composition and genetic background as specified in the table. One asterisk indicates significance at the .05 level and two at the .01 level*

Genetic background	Age and sex	Correlations between results of standard-GTT and cortisone-GTT at:							
		0 hours		1 hour		2 hours		3 hours	
		N	$r \pm \sigma_r$	N	$r + \sigma_r$	N	$r \pm \sigma_r$	N	$r \pm \sigma_r$
Controls	10–29, M, F	83	**0.316	84	**0.355	84	**0.338	81	**0.489
Offspring of conjugal diabetic parents.	10–29, M, F	20	0.389	20	*0.544	20	*0.553	18	0.354
Offspring of one diabetic and one normal parent.	10–29, M, F	107	**0.373	106	**0.490	105	**0.334	101	**0.533

The age span covered by these controls is unfortunately limited. However, through the kindness of Dr. C. J. Tupper, we have had access to the recent results of the periodic health appraisal of the senior staff of the University of Michigan. As one aspect of this evaluation, each individual received a screening, two point (0 and 2 hour) glucose tolerance test. In recent years (subsequent to the ex-

perience described on page 106), the patients have been asked to adhere to a preparatory diet. After the fasting blood specimen had been drawn, a standard test meal of 100 gm of glucose in lemon-flavored water was administered. The findings regarding the 2-hour blood sugar in those individuals with a negative family history for diabetes mellitus are given in table 4. Individuals thought to have abnormally high values at the 2-hour determination subsequently received the standard GTT, preceded by a prep diet. At the time of the evaluation, a medical history was obtained from each person by an internist.

TABLE 4.—*The blood sugar level 2 hours after 100 gms of glucose in those University of Michigan employees undergoing a routine health appraisal who gave a negative family history for diabetes mellitus. Subjects were predominantly of the male sex*

	Age				
	20–29	30–39	40–49	50–59	60–
$\bar{x}\pm\sigma_{\bar{x}}$	80.5±3.9	86.1±1.3	91.1±2.1	94.4±4.6	107.2±5.6
σ	18.0	20.7	22.1	23.9	26.7
N	21	267	107	27	23
No.>120 mgm %	0	18(6.7)	9(8.4)	3(11.1)	9(39.1)

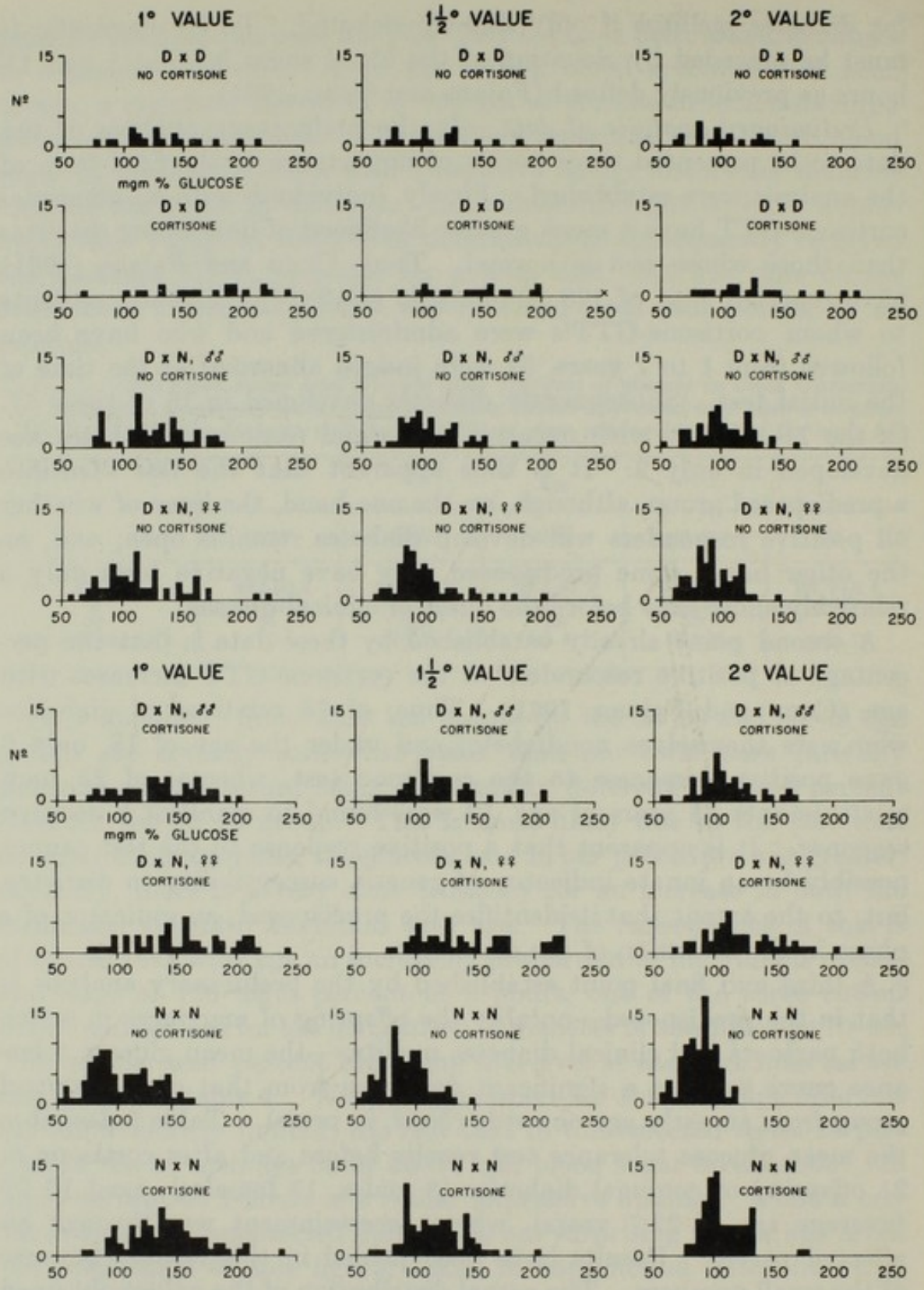
It is apparent, first, that the values for the 20–29 and 30–39 age groups are actually somewhat lower than our own more carefully screened control series, being significantly different at the 1 percent level for the 20–29 group. This is most likely due to the fact that on a per kilogram basis, the glucose load in our procedure was greater. Secondly, there is a very clear tendency for an increase in both the mean and standard deviation with age. The consequence of this is of course that with age an increasing proportion of individuals exceed the value of 120 mgm percent at 2 hours, one of the three cut-off points earlier used for the delineation of diabetes in a sample age 10–29. The number and percent exceeding this level at each age interval are also given in table 4. While the numbers for the older age intervals are unfortunately limited, the fact that in this selected series 24 percent of the 50 men age 50 or above had blood sugar levels about 120 mgm percent at 2-hours is a rather impressive finding. While a certain amount of undetected diabetes is not surprising at this age level, the magnitude of this figure, taken in conjunction with the other "control" figures presented earlier, not only raises some questions concerning the diagnostic standards for diabetes in the elderly, but, as we shall see, some especially troublesome issues for the interpretation of family studies. Using a different testing procedure (cortisone glucose tolerance test) we have previously suggested that diagnostic criteria may have to be modified with advancing age (Conn and Fajans, 1961). For clinical purposes, it needs to be pointed out that although the 2-hour post-glucose blood sugar level is an excellent screening procedure, it cannot be employed as a diagnostic procedure

for diabetes mellitus if only slightly elevated. To be diagnostic it must be preceded by elevation of the blood sugar level at 1 and 1½ hours as previously defined (Fajans and Conn, 1961).

Preliminary analysis of data.—In the preliminary analysis of the data to be presented three points of importance to the final form of the analysis were established. Firstly, individuals with an abnormal cortisone-GTT have a much greater likelihood of developing diabetes than those whose test is normal. Thus, Conn and Fajans (1961) have reported that of 128 nondiabetic relatives of diabetic patients to whom cortisone-GTT's were administered and who have been followed from 1 to 7 years, 57 were judged abnormal at the time of the initial test. Subsequently diabetes developed in 15 of these 57. Of the 71 subjects with negative or normal responses, diabetes has developed in only 2. It is thus apparent that the test identifies a predisposed group, although, on the one hand, the issue of whether all positive responders will develop diabetes remains open, and, on the other hand, some predisposed may have negative tests only a relatively short time before the onset of clinical disease.

A second point already established by these data is that the percentage of positive responders to the cortisone-GTT increases with age (Conn and Fajans, 1961). Thus, of 78 relatives of diabetics who were themselves nondiabetic and under the age of 18, only 6 gave positive responses to the cortisone test, whereas of 25 such relatives over 48 years of age, 12 were found to manifest a positive response. It is apparent that a positive response to the test cannot possibly be an innate indicator of genetic susceptibility to diabetes, but, to the extent that it identifies the predisposed, an indicator of a phase in the evolution of the disease.

A third and final point established by the preliminary analysis is that in the predisposed—notably the offspring of marriages in which both partners had clinical diabetes mellitus—the mean glucose tolerance curve exhibits a significant deviation from that of the control group from an early age onwards (Neel, in press). Table 5 describes the mean glucose tolerance test results before and after cortisone in 21 offspring of conjugal diabetics (8 males, 13 females), aged 10–29 (average age of 22.7 years), where ascertainment was through an affected parent. Results have been pooled in this fashion because of the small numbers. The actual distribution of the values obtained at 1, 1½, and 2 hours is shown in Fig. 1. In addition to the 21 offspring tested, there were in these families 8 other children in this age group who were not tested, none of whom were reported to have clinical diabetes. As before, correlation coefficients have been computed for four of the seven reference points before and after cortisone. These correlations are given in table 3. Again we see moderate correlations, possibly a little greater than seen in the controls, although the small numbers available make positive conclusions impossible.



FIGURES 1 a and b.—The distribution of the 1-, 1½-, and 2-hour blood glucose values on a standard- and cortisone-glucose tolerance test, in three classes of individuals: Offspring of conjugal diabetics, offspring of one diabetic and one normal parent (ascertained through diabetic parent), and controls. All individuals age 10–29. Further explanation in text.

Figs. 2 and 3 contrast the mean curves in this group with those observed in the controls. Table 6 gives a point-by-point comparison of the two curves. The D × D offspring consistently yield higher

TABLE 5.—*The results of standard and cortisone-GTT's in 21 offspring of conjugal diabetics of both sexes and aged 10–29. Mean age of the subjects was 22.7 years*

Type of test	Value $\bar{x}$	Hours after test meal						
		0	½	1	1½	2	2½	3
Standard	$\bar{x}\pm\sigma_{\bar{x}}$	86.5±1.6	132.0±5.3	132.1±7.6	113.0±8.2	103.2±5.9	101.0±6.2	93.8±4.4
	σ	7.1	24.3	34.7	37.5	27.0	28.5	20.3
	N	21	21	21	21	21	21	21
Cortisone	$\bar{x}\pm\sigma_{\bar{x}}$	99.0±2.2	160.1±5.7	164.4±10.1	148.2±10.6	128.6±8.4	120.8±7.3	109.2±6.2
	σ	10.0	25.6	45.0	46.1	37.7	31.9	26.4
	N	20	20	20	19	20	19	18

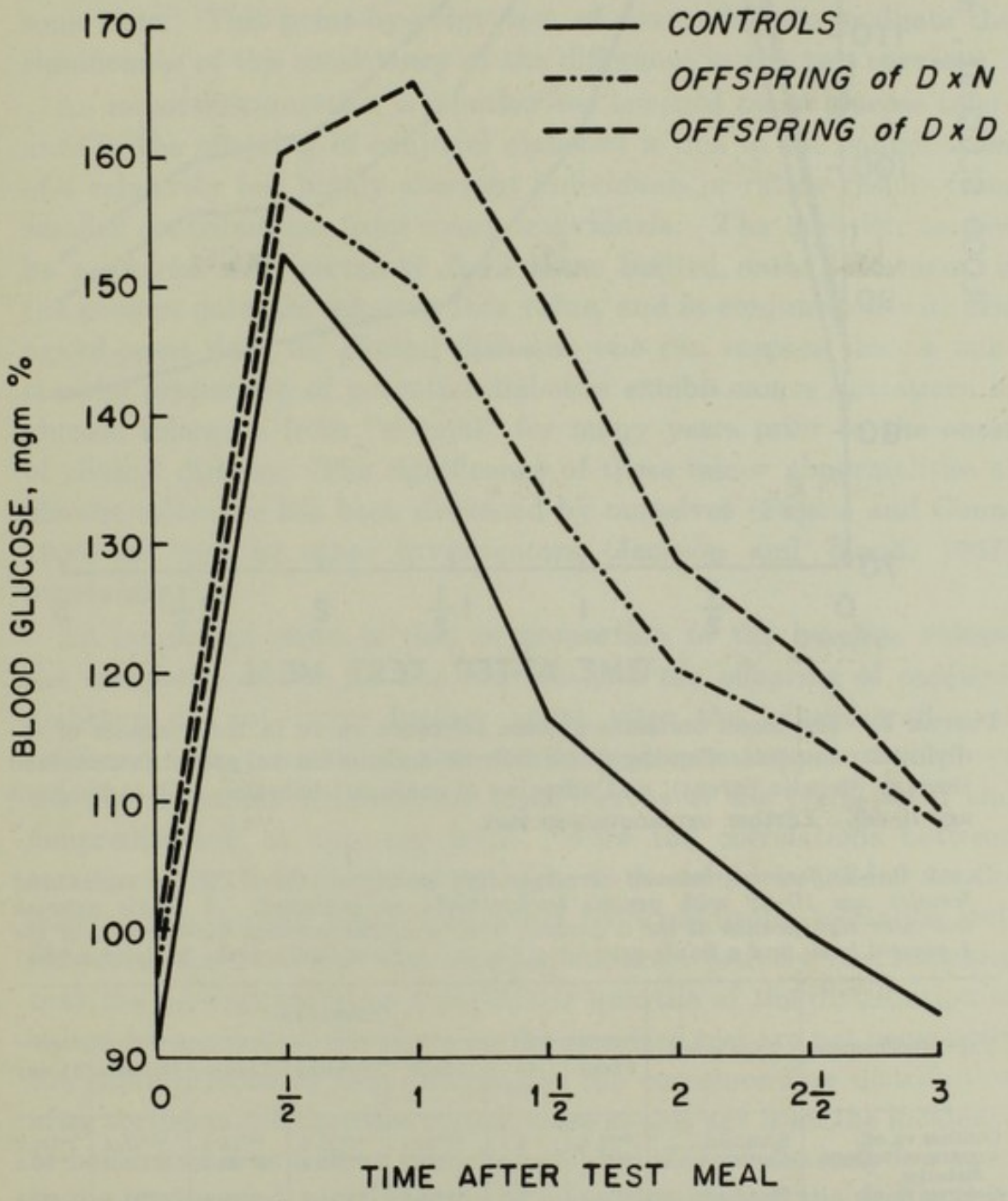


FIGURE 2.—The mean standard glucose tolerance curve in three classes of individuals: controls, offspring of one diabetic and one normal parent (ascertained through diabetic parent), and offspring of conjugal diabetics. All individuals age 10–29. Further explanation in text.

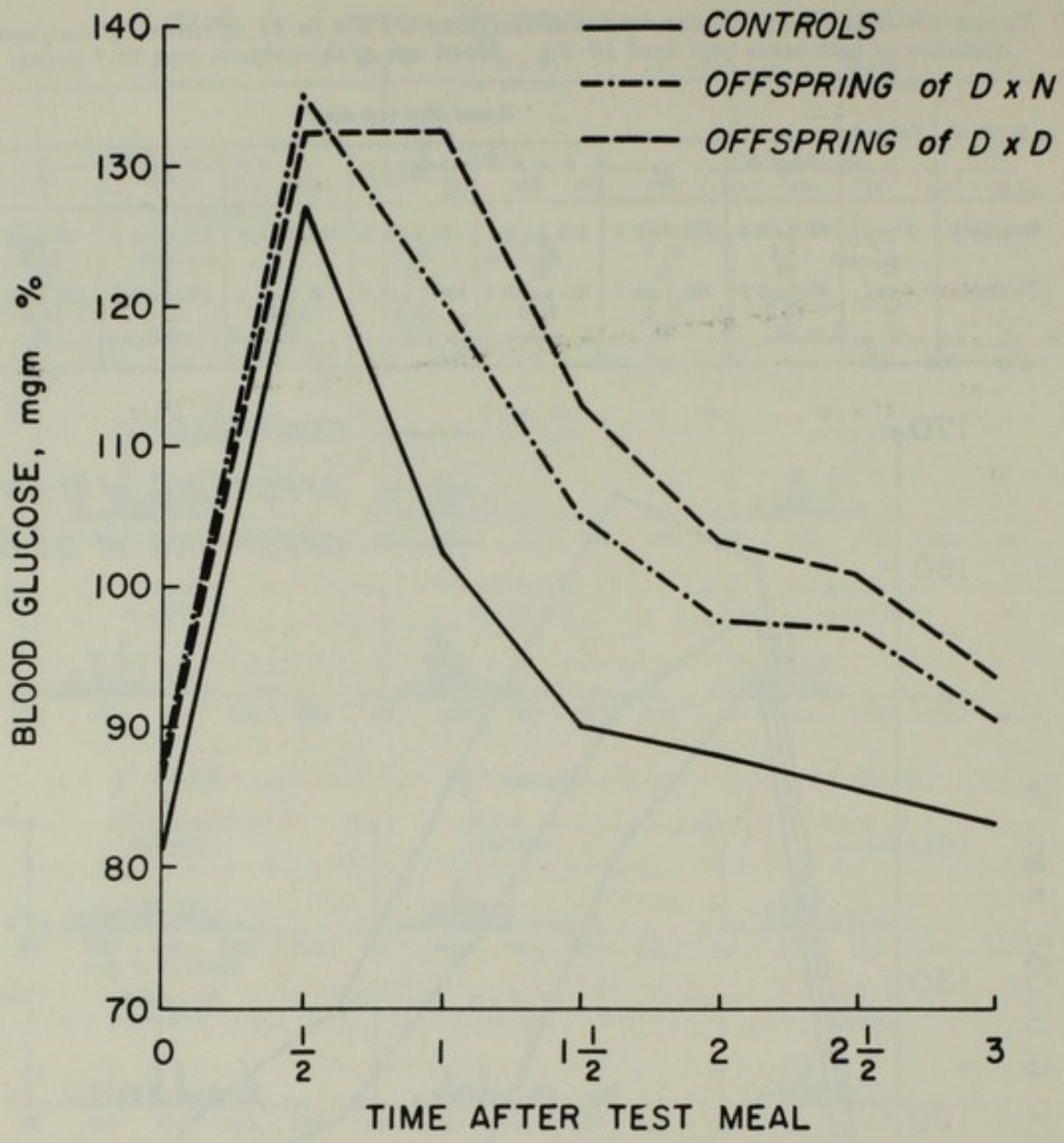


FIGURE 3.—The mean cortisone-glucose tolerance curve in three classes of individuals: controls, offspring of one diabetic and one normal parent (ascertained through diabetic parent), and offspring of conjugal diabetics. All individuals age 10–29. Further explanation in text.

TABLE 6.—*Differences between corresponding points on the GTT for males and females age 10–29 with genetic backgrounds as indicated. A single asterisk indicates significance at the 5 percent level, a double asterisk significance at the 1 percent level, and a triple asterisk at the 0.1 percent level, on the basis of a t-test*

Type of comparison	Type of test	Difference at:						
		0 hour	½ hour	1 hour	1½ hours	2 hours	2½ hours	3 hours
Controls vs. offspring of conjugal diabetics.	Standard	***5.0	4.1	***29.1	***22.9	***15.2	***15.4	**10.7
	Cortisone	***7.7	7.8	**25.9	***32.0	***22.3	***21.3	**16.1
Controls vs. offspring of D × N marriages.	Standard	***5.4	*7.1	***16.8	***14.8	***9.4	***11.5	**7.3
	Cortisone	***5.9	4.9	10.4	**17.1	**13.5	***15.1	***14.9
Offspring of conjugal diabetics vs offspring of D × N.	Standard	−0.4	−0.3	12.3	8.1	5.8	4.0	3.5
	Cortisone	1.9	2.9	15.5	14.9	8.8	6.2	1.2

mean values than the controls. However, because of the larger standard deviations for the mean values of the D × D offspring, the findings in this latter group overlap almost completely the performance of the control individuals. F tests reveal that the variance of the D × D offspring is significantly greater than that of the controls for 10 of the 14 possible comparisons. Under these circumstances, *t* tests comparing mean values on the glucose tolerance curves of controls and D × D offspring must be applied and interpreted with caution. These *t* tests reveal apparently significant differences in the means for six of the seven points for both the standard and cortisone tests. This point-by-point test of course fails to evaluate the significance of the consistency of the difference in the two curves.

An important question is whether the lessened mean glucose tolerance in the offspring of conjugal diabetics is due to the contribution of a relatively few highly aberrant individuals or rather results from smaller contributions from many individuals. The question cannot be answered with certainty from these limited data. However, if the present data are taken at face value, and in conjunction with the age-of-onset data for clinical diabetes, one can suspect that a substantial proportion of potential diabetics exhibit minor deviations in glucose tolerance from "normal" for many years prior to the onset of clinical disease. The significance of these minor abnormalities of glucose tolerance has been discussed by ourselves (Fajans and Conn, 1959) as well as other investigators (Jackson and Woolf, 1957) previously.

An important point is that in proportion to the baseline values, the composite curves for the controls and the offspring of conjugal diabetics do not move further apart after the administration of cortisone. On the average, then, cortisone does not appear to improve the discrimination between the mean curves of the predisposed and nonpredisposed at this age level. Were the correlations between corresponding individual values high, little would be added to the information on any given individual by a cortisone tolerance test. However, the fact that the correlations are relatively low indicates that the persons found in a particular quartile of the distribution of values for a point on the curve on the standard test are not necessarily the same persons in this quartile for the corresponding distribution after cortisone. Otherwise stated, since at this age level the incidence of both clinical as well as "subclinical" diabetes is relatively low even in the predisposed, a comparison of *mean* responses of the predisposed and nonpredisposed may hide a few important individuals in the former group in whom the cortisone-glucose tolerance test may make an important distinction.

735-712 O—65——9

We come at last to our test of hypothesis. We have derived, on the one hand, a mean GTT, with and without cortisone, for a group of young normal individuals, and similar curves for a group whose expectation of developing the disease would seem to approach 100 percent. We have also established standards for a qualitative judgment of the characteristics of a GTT. We now propose three approaches to the question of the genetic basis of diabetes mellitus, two quantitative and one qualitative. On the quantitative side, the mean type of glucose tolerance curve encountered in individuals age 10–29 who are the offspring of diabetic × normal (D × N) marriages will be examined, to determine whether the characteristics of this curve are consistent with the concept that all diabetes mellitus is due to homozygosity for the same recessive gene, i.e., of genotype *dd*. This examination will consist of 1) a study of the mean curve and its position relative to the two curves already established, to determine the proportionate representation in the offspring of D × N marriages of potentially diabetic and normal children; and 2) an examination of the distribution of individual values at specified reference points in the GTT, in an effort to distinguish between unimodality and bimodality. Then, on the qualitative side, 3) data will be presented in which the types of GTT encountered in older (40–59) offspring of D × N matings have been scored according to the previously established definitions.

It will be noted that we do not propose to use data on the 30–39 age group. This is felt to be a "transitional" group, which neither lends itself to the quantitative approach appropriate to the younger offspring nor the qualitative approach appropriate to the older. In the past, investigators have tended to consider all the children of a given mating type together. It is our contention that by concentrating on the age extremes, and applying an appropriate (different) technique to each extreme, the picture may be clarified.

In tables 7 and 8 are given the mean glucose tolerance curves, before and after cortisone, of 115 individuals aged 10–29 who are the offspring of diabetic × normal marriages, ascertained through the diabetic parent. The data are subdivided according to the decade and sex of the subjects. Excluded from the calculation are the results in eight individuals found grossly diabetic at the time of the glucose tolerance test. The findings in these eight persons are given in table 9. Although excluded from the derivation of the curves, they will enter into the later calculations. The rationale for exclusion is that at this level of abnormality second order effects begin to appear which render these individuals not *in pari materia* with the others. In addition, there were 3 siblings in this age range known to be diabetic

TABLE 7.—*Mean glucose tolerance curves, by decade and sex, of individuals with one diabetic and one normal parent, ascertained through the diabetic parent. The significance of all differences shown has been examined by a t-test, with one, two, and three asterisks indicating significance at the .05, .01, and .001 levels respectively*

Age	Sex	Value	Hours after test meal						
			0	½	1	1½	2	2½	3
10–19	M	$\bar{x}\pm\sigma_{\bar{x}}$	90.3±1.5	141.4±4.2	120.2±5.3	104.8±5.4	102.9±4.0	95.1±3.2	90.6±3.8
		σ	8.4	24.0	30.5	30.4	22.7	17.9	20.7
		N	33	33	33	31	32	31	30
20–29	M	$\bar{x}\pm\sigma_{\bar{x}}$	87.0±1.9	143.8±5.4	132.5±5.6	111.7±5.2	98.7±4.9	103.8±4.7	91.2±5.4
		σ	8.3	24.3	25.2	23.5	21.9	21.2	23.9
		N	20	20	20	20	20	20	20
10–19	F	$\bar{x}\pm\sigma_{\bar{x}}$	87.6±1.4	130.3±5.9	117.5±6.9	104.3±6.4	97.1±4.0	98.3±4.8	94.6±5.0
		σ	7.0	29.4	35.2	31.9	20.2	23.4	23.7
		N	26	25	26	25	26	24	23
20–29	F	$\bar{x}\pm\sigma_{\bar{x}}$	83.3±1.2	127.2±4.0	113.9±5.5	101.6±4.0	92.2±3.3	94.4±3.4	86.9±2.5
		σ	7.4	24.0	32.8	23.5	19.8	20.3	14.7
		N	36	36	35	36	36	36	36
20–29	M–F	Diff.	3.7	16.6*	18.6*	10.1	6.5	9.4	4.3
10–29	M, F†	$\bar{x}\pm\sigma_{\bar{x}}$	86.9±0.8	134.9±2.4	119.8±3.0	104.9±2.6	97.4±2.0	97.1±2.0	90.3±1.9
		σ	8.2	26.0	31.8	27.4	21.3	20.6	20.2
		N	115	114	114	112	114	111	109

†Mean age of the subjects was 19.8 years.

TABLE 8.—*Mean cortisone-glucose tolerance curves, by decade and sex, of individuals with one diabetic and one normal parent, ascertained through the diabetic parent. The significance of all differences shown has been examined by a t-test, with one, two, and three asterisks indicating significance at the .05, .01, and .001 levels respectively*

Age	Sex	Value	Hours after test meal						
			0	½	1	1½	2	2½	3
10–19	M	$\bar{x}\pm\sigma_{\bar{x}}$	95.0±1.4	150.9±4.6	124.7±5.3	111.5±3.7	105.8±3.6	100.5±2.9	98.0±4.5
		σ	8.1	26.3	30.0	21.1	20.4	16.3	24.5
		N	32	31	32	31	32	31	30
20–29	M	$\bar{x}\pm\sigma_{\bar{x}}$	94.2±2.4	155.3±6.1	158.7±8.5	136.4±9.5	115.4±7.5	114.4±8.2	105.2±5.4
		σ	10.2	25.8	35.9	40.4	31.0	34.6	23.0
		N	18	18	18	18	17	18	18
10–19	F	$\bar{x}\pm\sigma_{\bar{x}}$	100.9±2.8	156.7±7.8	156.0±10.8	140.6±9.9	130.7±9.6	125.1±9.0	115.7±10.5
		σ	14.3	39.2	55.1	50.6	48.7	46.1	53.7
		N	26	25	26	26	26	26	26
20–29	F	$\bar{x}\pm\sigma_{\bar{x}}$	97.9±1.8	165.0±6.2	165.4±9.8	147.1±9.0	127.5±6.8	120.3±5.7	113.0±4.5
		σ	10.1	34.4	54.4	49.8	37.6	31.9	24.9
		N	31	31	31	31	31	31	30
20–29	M–F	Diff.	−3.7	−9.7	−6.7	−10.7	−12.2	−5.9	−7.8
10–29	M, F	$\bar{x}\pm\sigma_{\bar{x}}$	97.1±1.1	157.2±3.1	149.8±4.6	133.3±4.3	119.8±3.6	114.7±3.3	108.0±3.4
		σ	10.9	32.2	48.0	43.8	36.7	34.1	34.4
		N	107	106	107	106	106	106	104

TABLE 9.—*The findings in 7 individuals with grossly abnormal, standard GTT's, and for this reason excluded from the calculations of tables 7 and 8*

Individual	Age	Sex	Hours after test meal						
			0	½	1	1½	2	2½	3
MM–8170	11	M	272	368	437	456	452	473	444
GW–5577	29	M	147	264	337	357	355	287	229
TW–5577	27	M	210	328	405	455	451	435	392
MW–5577	25	M	113	236	274	267	259	216	175
DW–5577	11	M	89	175	209	221	227	210	180
TH–8156	22	M	124	190	253	193	166	198	188
IH–8156	17	F	153	238	288	323	323	316	312
PT–8232	28	F	114	181	243	253	238	227	184

and not tested, and 35 who were not known to be diabetic who could not be tested.

Again, before deriving a composite curve we must examine the data for possible age and sex differences. Firstly, for the 20–29 age group, males (as was true for the controls) exhibit higher values than females, greater in absolute value than in the control material, but significant at the 5 percent level for only two of the seven comparisons. Again (as observed in the controls) we note the reversal of the sign of the difference after cortisone; again none of the differences is significant. By this criterion, then, the cortisone test reduces the difference between sexes. Secondly, inspection of the data again raises the possibility of age effects. All seven mean values in the standard curve of the 20–29 female group are *lower* than in the 10–19 group, although none of these differences is significant. More striking is the difference between the cortisone-GTT's of 10–19 and 20–29 year old males, significant for the 1- and 1½-hour points. Study of the individual curves suggests that this age effect is due primarily to a disproportionately low response of the 10–19 M group to cortisone, as contrasted to the 20–29 M group, or to the controls. The accentuated response of females to cortisone may be related to a potentiation of the biological activity of this corticosteroid by estrogens (Nelson, Tanney, Mestman, Gieschen, and Wilson, 1963).

In the strict sense, the data do not permit combining the various age and sex groups. Since, however, as we shall see, we view this analysis as very approximate, we have in tables 7 and 8 for subsequent purposes combined the data for both sexes and both age intervals into single curves, as was done for the controls (tables 1 and 2) and the offspring of conjugal diabetics (table 5). Table 6 reveals that most points on the resulting two curves differ significantly from the corresponding points on the two control curves. On the other hand, the corresponding differences from the curves based on the offspring of conjugal diabetics are not significant, and, in fact, the D × N offspring exceed the D × D offspring for the first two points.

We are now ready for the first *approximate* test of hypothesis. If the marriage of diabetics and normals produces two types of individuals, we might examine the data for evidence for this at those time intervals in the GTT which provide the best test for the ability of an individual to return his blood sugar to normal levels. These are probably the 1-, 1½-, and 2-hour samples. In figure 1 are shown the findings, for these time points, in the following groups: 1 and 2) offspring of D × D, 10–29, sexes combined, before and after cortisone; 3–6) offspring of D × N, 10–29, sexes separately (because of the more marked sex difference), before and after cortisone; 7 and 8) offspring of N × N, 10–29, sexes combined, before and after cortisone. By presenting the sexes separately one avoids the danger of obscuring a bimodality because the curves being combined are somewhat out

of phase. The greater variability in the offspring of conjugal diabetics than in the controls, already expressed in the standard deviations of tables 1, 2, and 5, is evident. The individual blood sugar determinations suggest that at this age level, some of the offspring of conjugal diabetics are still essentially normal in their glucose metabolism, with all variations between this state and florid diabetes. Under these circumstances, one would not expect to demonstrate two clearly modal values among the offspring of diabetic × normal marriages, and indeed only one mode is apparent. This approach, then, has failed to provide a satisfactory test of a one-recessive-gene hypothesis.

The second approach concerns itself with a comparison of the mean values observed in the three types of matings. If diabetes is inherited as a recessive trait, then one can readily calculate the proportion of affected offspring to be expected from a sample of D × N marriages, *provided* there are data on 1) the frequency in the population of the assumed homozygous recessive genotype and 2) the relative frequency among the phenotypically normal parents of *DD*, *Dd*, and *dd*, i.e., of homozygous normal, carriers, and phenotypically normal individuals who may be prediabetics or subclinical or latent diabetics. Unfortunately, as we have seen, neither point 1) nor 2) can be established with any degree of accuracy. A "middle-of-the-road" calculation would proceed as follows: Let us assume a 0.10 frequency at birth of the diabetic predisposition, i.e., *dd*, which implies a frequency of the *d* gene of 0.317. Let us further assume that among the partners of the D × N marriages, 0.50 of those predisposed have already developed diabetes. The frequency of *DD*, *Dd*, and *dd* among the phenotypically normal parents then becomes $\frac{0.683^2}{1.00-0.05}=.4910$, $\frac{2(0.683)(0.317)}{1.00-0.05}=.4559$, and $\frac{0.10-0.05}{1.00-0.05}=.0526$. From the *Dd* × *dd* marriages half of the children will be *dd;* assuming equal fertility, this will amount to 0.23 of all children from D × N marriages. From the *dd* × *dd* marriages all the children will be involved; again assuming equal fertility this will amount to 0.05. Total expectation of *dd* offspring from D × N marriages is thus 0.28. If we prefer to assume a frequency of the hypothetical *dd* genotype as low as 0.07, and that none of the phenotypically normal partners in the D × N cross are potential diabetics, then the proportion of affected to be expected is approximately 0.21, whereas if we assume a frequency of the predisposition of 0.20 and that the phenotypically normal parents include the diabetic genotype in the proportion of potential diabetics present in the population at birth, then the figure is approximately 0.45. Quite obviously, until a better estimate is available of the frequency of the diabetic predisposition (genotype) in the population, tests of hypothesis can only be approximate.

Figures 2 and 3 depict the position of the mean glucose tolerance curve of the D × N offspring age 10–29 (average age 19.8 years) relative to the controls and the offspring of conjugal diabetics. From the tables thus far presented, we can, by utilizing corresponding time points (1, 1½, and 2 hours) on the appropriate curves, readily calculate the number and proportion in which ultimately diabetes would be expected to develop among the offspring of D × N marriages, assuming the potential diabetics behave as do the offspring of D × D marriages, and the normals as the offspring of N × N marriages. For example, for the standard GTT at 1°, $\frac{132.1y+103.0(114-y)}{114}=119.8$, where y is the number of potential diabetics; y=65.8 and $P_y=0.58$. The results of these six calculations, *which have an ill-defined but relatively enormous error*, are given in table 10. The (unweighted) average proportion of potential diabetics among 115 persons tested by one or both types of GTT at six different time points is estimated to be 0.57. In addition, these 115 persons had, within that 10–29 year interval, 3 known diabetic siblings, 8 with newly recognized latent diabetes (see table 9), and 35 untested sibs. Assuming that half of the untested 35 would be found to be abnormal, then the proportion of individuals with the diabetic predisposition in this group would be estimated at $\frac{0.57\times115+3+8+0.5\times35}{115+3+8+35}=\frac{94.1}{161}=0.58$. This is approximately twice as high a proportion as our gene frequency analysis led us to expect on the assumption of simple recessive inheritance and a frequency of 0.10 for the diabetic predisposition. Only by assigning frequencies to the homozygous recessive genotype of the order of 0.20–0.25 can we reconcile the findings with simple recessive inheritance. However, it must be borne in mind how rough this approach really is, as well as the indeterminate but undoubtedly very large error. For instance, in addition to the sampling errors, differences in the age and sex composition of the three samples used in the estimate might introduce errors in the calculation. In this connection, it should be noted that the mean ages of the control subjects, offspring of D × D, and offspring of D × N marriages in the 10–29 age interval, are respectively, 24.7, 22.7, and 19.8 years. Furthermore, as is apparent from the appropriate tables, females predominate in the D × D sample but males in the controls. While no correction for these differences seems practical, here is a further indication of the need for caution in the literal interpretation of these findings. Finally, it must be recalled that neither the N × N or D × D curves are "pure," in that the former may contain some potential diabetics and the latter some non-potential diabetics. This dilution of the two curves would tend to bring them closer together; whether it introduces an appreciable error depends on the degree to which it occurs in each of the curves.

TABLE 10.—*The proportion of diabetics to be expected from the offspring of D × N marriages, as computed before and after cortisone and at three different points on the glucose tolerance curve*

	1°		1½°		2°	
	Number	Proportion	Number	Proportion	Number	Proportion
No cortisone	65.8	0.58	72.4	0.65	70.5	0.62
Cortisone	44.3	0.47	56.6	0.53	64.2	0.61

There is a second approach to estimating the proportion to be affected in a D × N mating. Among the offspring of D × N marriages in the families studied there are 79 in the 40–59 age range. Of these, 8 have previously diagnosed diabetes, 10 latent diabetes detected during this study, 14 abnormal cortisone-GTT's, 21 normal GTT's (7 of these did not have the cortisone test), and 26 are not known to have diabetes but were untested. Treating the first three types as all of the diabetic genotype, then the minimum proportion would be 30/79, or 0.38. If roughly a half of those untested have the diabetic predisposition (on the basis of the observations both on this group and the 10–29 age group), then the proportion rises to 43/79, or 0.54. Again, then, we would appear to have an excess of diabetically predisposed over expectation on the basis of a single recessive gene, given in assumptions. However, it must be borne in mind that the diagnostic standards for both the standard and cortisone-GTT applied here were derived from a younger group; the data of the Faculty Health Appraisal described earlier suggest these standards may be somewhat inadequate at this age level, although it should be noted that taken at face value the "age effect" on glucose metabolism is much more marked in the seventh and subsequent decades than in the 49–59 interval. This fact might lead to an overestimate of the proportion with the diabetic genotype. On the other hand, some of those with normal GTT's may yet develop abnormality—this would result in an underestimate.

DISCUSSION

The approach to the genetic basis of diabetes presented in this paper is quite laborious; the data collected to date are adequate only for very tentative conclusions. While there is no doubt of the familial nature of the disease, we have not, in the sense of adherence to genetic ratios, established that this is on a genetic basis. Rather, we believe it to be genetic because we have no other explanation for the familial constellations of the disease. In this respect, the case for a genetic basis for diabetes is no weaker than for essential hypertension or hyperuricemia. Under these circumstances, the clearest evidence for

a genetic etiology comes from the data of which there is need for much more, on twins, involving not only pairs, one or both of whom are diabetic, but elderly pairs, neither known to be diabetic. However, if we assume a genetic basis and proceed to test the hypothesis of recessive inheritance, there appears to be an excess of diabetically predisposed among the offspring of D × N marriages over expectation on the hypothesis that the disease is due to homozygosity for recessive genes occurring at a single genetic locus, unless one is willing to assign to the diabetic genotype much higher frequencies than have been acceptable to date, of the order of 0.20–0.25. Since if recessive genes at several loci were involved expectation would be even lower, the findings are thus not consistent with this hypothesis either. Accordingly, the strongest piece of evidence for the recessive inheritance of diabetes is the report that all the offspring of conjugal diabetics appear destined to develop the disease. However, as we have pointed out, Post's report as amended in this paper, does not exclude the hypothesis that only 75 percent of these offspring will develop diabetes. Other objections to the recessive gene hypothesis have been brought forward by Allan (quoted in Steinberg, 1961) and Simpson (1962).

What alternative hypothesis can be suggested? The proportions given are not inconsistent with a hypothesis of simple dominant inheritance, although a precise test is again impossible because of uncertainty regarding the frequency of the diabetic genotype in the population. However, Steinberg (1959, 1961) and Simpson (1962) have advanced what appear to be rather serious objections to the hypothesis of simple dominant inheritance or to a hypothesis that regards severe, juvenile diabetes as due to homozygosity for a gene which, when heterozygous, produces in some individuals a milder, "adult-type" diabetes. We would appear, then, to be forced either to somewhat more complex, multigenic hypotheses, involving such possibilities as a "principal gene" with modifiers, or genes at several different loci with approximately additive effect (cf. Lamy, Frézal, and Rey, 1961; Simpson, 1962) or to regarding the disease as heterogeneous, the different entities exhibiting different inheritance patterns, a viewpoint adumbrated in Harris (1949, 1950). The evolution of genetic thought concerning diabetes would thus appear to parallel that concerning essential hypertension and hyperuricemia, for some time regarded as single gene traits but now increasingly being considered as possibly due to the interaction of genes at several different loci (Hamilton, Pickering, Fraser Roberts, and Sowry, 1954; Hauge and Harvald, 1955).

A multi-genic hypothesis can take a variety of forms, too many to be explored in detail at this time. Progress in refining this hypothesis is inextricably bound to advances in our understanding of the pathophysiology of the disease. It is also bound to a different approach,

so that instead of studying special families, the investigator examines a suitable population sample of families, to the data on which he can apply the techniques of regression, correlation, and variance analysis. The problem of age correction will remain controversial until better longitudinal data are available. However, one point requires comment: Under most such models, while a significant parent-offspring correlation is of course to be expected, one does not expect all of the children of affected parents to be involved. If, however, the changing environment is resulting in a progressively lower cut-off point in very recent generations on a quasi-continuous curve of diabetic predisposition, then one might expect a very high proportion of the children of clinical diabetics to develop the disease themselves. But if this is the explanation of the high proportion of affected children when both parents are affected, it implies that the frequency of diabetes is increasing rather markedly. Unfortunately, our prevalence figures have until recently been so poor that this hypothesis cannot be tested, but should it be proved correct some 20 years hence, the genetic approach will have provided a clear lead.

The high frequency of the genotype predisposing to so serious a disease as diabetes mellitus presents a genetic riddle. Why has natural selection not reduced the frequency of this genotype to a much lower level? Recently one of us (Neel, 1962) has suggested that the genotype may be involved in the ability to mobilize insulin readily following the proper stimulus. It is further suggested that a genotype which might have been optimal during man's hunting and gathering days (and then only rarely resulted in diabetes mellitus) may result in an overproduction of insulin under present conditions. This overproduction is countered by the development of insulin antagonists, which eventually come to dominate the picture, with resulting diabetes. This hypothesis of the role of the diabetic genotype is to a considerable extent independent of hypotheses concerning the precise genetic mechanism responsible for the disease. That is to say, the hypothesis can as well be applied to a one-gene as to a multi-gene concept of the diabetic predisposition.

The genetic approach can often result in a significant clarification of a heterogeneous disease entity. A notable example of this concerns the sickle cell diseases (Neel, 1952). Equally often, however, significant progress on the genetic front must await increased data concerning basic pathophysiology. This would seem to be the case with respect to diabetes. The evidence suggests that the abnormalities of the GTT of the diabetically predisposed are paralleled by, and presumably the result of, abnormal relationships between circulating insulin and insulin antagonists (references in Neel, 1962). Disease of the basement membrane of the blood vessels appears also to be present from an early stage (Marble, in press). Obviously longitudinal studies of the diabetically predisposed, beginning at a very early age,

have much to offer, their difficulty notwithstanding. Such studies should lead us to the primary lesion, upon which sounder genetic investigations can be built.

SUMMARY

As a disease entity, diabetes mellitus presents many impediments to a precise genetic study. An investigation is described which attempts to meet some of these impediments by 1) the use of the glucose tolerance test and the cortisone-glucose tolerance test, and 2) the application of a quantitative approach to the data on the younger offspring of certain types of marriage but a qualitative approach to the older offspring of these same types of marriage. The data obtained once again confirm the familial nature of diabetes but in the strict and formal sense are not proof of a genetic predisposition although, by exclusion, this seems highly probable. It is shown that in the diabetically predisposed (offspring of conjugal diabetics) there are highly significant changes in the mean glucose tolerance curve in the 10–29 age interval. This fact can be utilized in the analysis of family data. Assuming a genetic basis, then the results of our analyses do not favor the hypothesis that diabetes is due to homozygosity for a single recessive gene, unless one is willing to assign a frequency of 0.20–0.25 to the recessive genotype. There appear to be valid reasons for discarding other simple genetic hypotheses. The data are consistent with a multigenic (multifactorial) hypothesis, although the apparently very high proportion of affected offspring from conjugal diabetics remains a troublesome point. This is explicable if the living conditions of the current generation have lowered the threshold of phenotypic expression of the diabetic genotype below that of the parental generation.

REFERENCES

1. Allan, W., 1933. Heredity in diabetes. *Ann. Intern. Med.* 6: 1272–1274.
2. Brandt, R. L., and Tupper, C. J., 1960. Medical appraisal of elderly county hospital patients. *Geriatrics* 15: 233–253.
3. Chesrow, E. J., and Bleger, J. M., 1955. Diabetes case finding among 1,000 patients over 60. *Geriatrics* 10: 479–486.
4. Cohen, A. M., 1963. Effect of environmental changes on prevalence of diabetes and of atherosclerosis in various ethnic groups in Israel. In *The Genetics of Migrant and Isolate Populations*, E. Goldschmidt, ed., pp. 127–130. Baltimore: Williams & Wilkins Co.
5. Conn, J. W., and Fajans, S. S., 1961. The prediabetic state. *Amer. J. Med.* 31: 839–850.
6. Fajans, S. S., and Conn, J. W., 1954. An approach to the prediction of diabetes mellitus by modification of the glucose tolerance test with cortisone. *Diabetes* 3: 296–304.

7. Fajans, S. S., and Conn, J. W., 1959. The early recognition of diabetes mellitus. *Ann. New York Acad. Sci.* 82: 208–218.
8. Fajans, S. S., and Conn, J. W., 1961. Comments on the cortisone-glucose tolerance test. *Diabetes* 10: 63–67.
9. Goldschmidt, E.; ed., 1963. *The Genetics of Migrant and Isolate Populations*, pp. xxi and 369. Baltimore: Williams & Wilkins Co.
10. Grunnet, J., 1957. Heredity in diabetes mellitus. *Opera ex Domo Biologiae Hereditariae Humanae Universitatis Hafniensis*, no. 39. Copenhagen: Munksgaard.
11. Hamilton, M., Pickering, G. W., Fraser Roberts, J. A., and Sowry, G. S. C., 1954. The etiology of essential hypertension. 4. The role of inheritance. *Clin. Sci.* 13: 273–304.
12. Harris, H., 1949. The incidence of parental consanguinity in diabetes mellitus. *Ann. Eugen.* 14: 293–300.
13. Harris, H., 1950. The familial distribution of diabetes mellitus: a study of the relatives of 1241 diabetic propositi. *Ann. Eugen.* 15: 95–119.
14. Hauge, M., and Harvald, B., 1955. Heredity in gout and hyperuricemia. *Acta Med. Scand.* 152: 247–257.
15. Hayner, N. S., Waterhouse, A., and Gordon, T., 1963. The 1-hour oral glucose tolerance test. *Health Statistics* (USPHS Publ. 1000), series 2, no. 3. Washington: GPO.
16. Jackson, W. P. U., and Woolf, N., 1957. Further studies in prediabetes. *Lancet* 1: 614–617.
17. Joslin, E. P., *et al.*, 1959. *Treatment of Diabetes Mellitus*, edition 10:798. Philadelphia: Lea & Febiger.
18. Lamy, M., Frézal, J., and Rey, J., 1961. Hérédité du diabète sucré. *Journées Ann. Diabèt. Hotel Dieu* 2: 5–15.
19. Levit, S. G., and Pessikova, L. N., 1934. The genetics of diabetes mellitus. *Proc. Maxim Gorky Medico-biological Research Institute* 3: 132–147. (Russian.)
20. Marble, A. Vascular changes in prediabetes. In *Conference on Small Blood Vessel Involvement in Diabetes Mellitus*, M. D. Siperstein, ed. (in press).
21. Neel, J. V., 1952. Perspectives in the genetics of sickle cell disease. *Blood* 7: 467–471.
22. Neel, J. V., 1962. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Amer. J. Hum. Genet.* 14: 353–362.
23. Neel, J. V. The genetic basis for diabetes mellitus. In *Conference on Small Blood Vessel Involvement in Diabetes Mellitus*, M. D. Siperstein, ed. (in press).
24. Nelson, D., Tanney, H., Mestman, G., Gieschen, V., and Wilson, L., 1963. Potentiation of the biologic effect of administered cortisol by estrogen treatment. *J. Clin. Endocrinol. Metab.* 23: 261–265.
25. Petri, L. M., McLaughlin, C. D., and Hodgins, T. E., 1954. Mass screening for lowered glucose tolerance. *Ann. Int. Med.* 40: 963–967.
26. Pincus, G., and White, P., 1933. On the inheritance of diabetes mellitus. I. An analysis of 675 family histories. *Am. J. Med. Sci.* 186: 1–14.
27. Pincus, G., and White, P., 1934a. On the inheritance of diabetes mellitus. II. Further analysis of family histories. *Am. J. Med. Sci.* 188: 159–168.
28. Pincus, G., and White, P., 1934b. On the inheritance of diabetes mellitus. III. The blood sugar values of the relatives of diabetics. *Am. J. Med. Sci.* 188: 782–790.
29. Post, R. H., 1962. An approach to the question, does all diabetes depend upon a single genetic locus? *Diabetes* 11: 56–65.
30. Randle, P. J., Garland, P. B., Hales, C. N., and Newsholme, E. A., 1963. The glucose fatty-acid cycle. Its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus. *Lancet* I: 785–789.

31. Rush, T., and Tupper, C. J., 1960. Two-hour postprandial glucose determinations in a periodic health appraisal program. *Geriatrics* 15: 630–636.
32. Simpson, N. E., 1962. The genetics of diabetes: a study of 233 families of juvenile diabetics. *Ann. Hum. Genet.* 26: 1–21.
33. Siperstein, M. D., ed. *Conference on Small Blood Vessel Involvement in Diabetes Mellitus.* (In press.)
34. Spiegelman, M., and Marks, H. H., 1946. Sex and age variations in the prevalence and onset of diabetes mellitus. *Am. J. Pub. Health 36:* 26–33.
35. Steinberg, A. G., 1959. The genetics of diabetes: a review. *Ann. York New Acad. Sci.* 82: 197–207.
36. Steinberg, A. G., 1961. Heredity in diabetes mellitus. *Diabetes* 10: 269–274.
37. Then Bergh, H., 1938. Die Erbbiologie des Diabetes Mellitus. Vorläufiges Ergebnis der Zwillingsuntersuchungen. *Arch. f. Rassen.-u. Gesellschaftsbiologie* 32: 289–340.
38. U.S. National Health Survey, 1960. Health Statistics: Diabetes reported in interviews, United States, July 1957–June 1959. *Pub. Health Ser. Publ.* no. 584–B21, pp. 1–22.
39. Vallance-Owen, J., and Lilley, M. D., 1961. Insulin antagonism in the plasma of obese diabetics and prediabetics. *Lancet* I: 806–807.
40. West, K. M., 1960. Response to cortisone in prediabetics. *Diabetes* 9: 379–385.

Coronary Artery Disease*

By Victor A. McKusick, M.D., *Professor of Medicine, Johns Hopkins University School of Medicine, Baltimore, Md.*

With regard to the relative importance of genetic and environmental factors in pathogenesis, it may be useful to think of all diseases of mankind as falling on a spectrum (fig. 1), toward one end of which genetic factors dominate and toward the other end environmental factors. Essentially all disease is to some extent genetic and to some extent environmental. Close to the genetic end lie conditions such as phenylketonuria and galactosemia, but they lie not at the extreme end because nongenetic influences can affect considerably the clinical, or phenotypic, features. Near the middle perhaps is diabetes mellitus and possibly not far beyond should be put hypertension and peptic ulcer. Toward the environmental end are infectious diseases, e.g., tuberculosis, but not at the extreme end if we can interpret the findings in twins to indicate a significant, genetically based difference in susceptibility to this disease. Where we should place atherosclerosis and specifically coronary artery disease, is the question we will be discussing.

There are distinctive characteristics of the diseases located toward one or the other end of the spectrum (fig. 1):

1. Those toward the "G" end tend to be rare; those toward the "E" end are common disorders, defining "common" as potentially present in at least 1 percent of persons at some stage in life.
2. Those toward the "G" end tend to be monogenic (i.e., monomeric) and relatively simple in their formal genetics. In those toward the "E" end the genetic contribution to pathogenesis tends to be complex and almost certainly polygenic. (I shall use the term *multifactorial* to indicate the operation of multiple factors both genetic and nongenetic and *polygenic* to indicate the operation of multiple genetic factors.)

*From the Division of Medical Genetics, Department of Medicine, Johns Hopkins University School of Medicine.

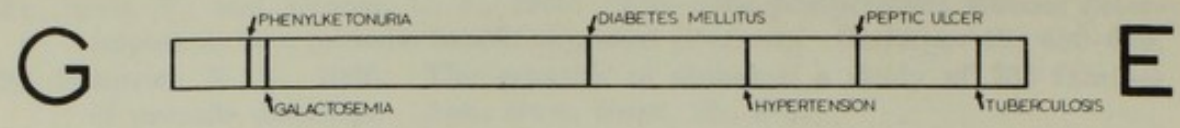


FIGURE 1.—Schematic representation of spectrum on which all disease falls with regard to the relative importance of genetic and environmental factors in pathogenesis.

In studying the genetics of disorders near one or the other end of the spectrum the questions one poses tend to be different. At the "G" end the questions concern 1) the details of the formal genetics (gene frequencies, mode of transmission, selection factors, mutation rate, etc.) and 2) the precise nature of the gene-determined basic defect which is assumed to be unitary. At the "E" end, the question tends to be, first and foremost: is there a significant genetic contribution to pathogenesis? Related to this question is a second: by what mechanisms do genetic factors contribute to the pathogenesis? (In principle the questions are not greatly different; the main difference lies in the level of analysis.)

The general approaches toward getting an answer to the question of the strength and mechanism of genetic factors in common diseases are of at least six types (Edwards, 1960; Edwards, 1963; Penrose, 1953):

1. Study of familial aggregation
2. Twin study
3. Interracial comparisons
4. Genetic study of component factors in pathogenesis
5. Blood group and other associations
6. Genetic study of animal analogs (or homologs).

Familial aggregation is essentially a *sine qua non* for any genetic hypothesis. Studies are of two general types depending on methods of selection of families: 1) proband studies in which index families are compared with control families, and 2) population-based studies such as the Tecumseh one (Francis, this symposium) in which familial clustering is tested for. The choice of controls is one source of difficulty in proband studies. Another is bias of ascertainment of the type that Dr. Crow (this symposium) discussed with regard to rare recessive traits. The selection biases in the latter situation have been recognized for a long time—since Weinberg. Haenzel (1959) and others before him have pointed out that in proband family studies of common disorders the same biases may operate. In elementary terms if a family with multiple cases of the disease has a greater probability of entering your series than does a family with only one case, then you cannot fail but find familial aggregation.

Familial aggregation does not, of course, prove the genetic basis of the trait in question because socio-environmental factors, as well as genes, vary from family to family. Determination of the proportion of offspring affected according to whether both, one or neither

parent is affected is a further test which is usefully applied to family data, although it too is not conclusive proof of the genetic hypothesis. The proportion of affected offspring is expected to be greatest with both parents affected, less with one parent affected, least with neither parent affected. It may be worthwhile to recall the story of pellagra.

In 1916, before pellagra was identified as a nutritional disease, it was thought to be due to a toxic agent, possibly infectious, susceptibility to which was strongly determined by genotype. In the studies of Davenport (1916) and Muncey (1916), the arguments for a prominent genetic factor in pellagra were 1) familial aggregation, 2) ethnic differences, specifically a lower frequency in Negroes, and 3) intrafamilial similarities in the clinical pattern of the disease. Although it apparently escaped their attention, evidence against the genetic factor was provided by the proportion of children affected according to whether both parents, one parent, or neither parent was affected (see table 1). A trait which is genetically determined to a significant extent should show a decreasing proportion of affected offspring according to whether both, one, or neither of the parents display the trait. Such is not the case in the data on pellagra.

TABLE 1.—*The proportion of pellagrous children according to parental mating type*
(Adapted from Muncey)

Number of parents pellagrous	Total number of children	Pellagrous offspring	
		Number	Percent
0	413	133	32.2
1	752	120	16.0
2	120	22	18.3
Total	1,285	375	29.2

It should be noted, however, that a positive "mating-type" test, although expected of a genetic trait, does not of itself prove the genetic factor. One can easily imagine a concentration of socio-environmental factors in families which would result in more "affected" children when more of the parents are "affected," and vice versa. Lilienfeld (1959) has shown that the "mating-type" test is positive for the trait "medical school attendance"; also a greater proportion of offspring are likely to attend university if the parents did than if the parents did not. Cigarette smoking in teenagers likewise shows a positive "mating-type" test (Horn *et al.*, 1959).

Twin studies are more specifically indicative of genetic factors as opposed to nongenetic ones. Dr. Harvald (this symposium) has outlined the difficulties of this approach. Of all of them, however, the twin method is perhaps the best we have for quantitating the role of genetic factors in common diseases.

Illustrating his point with an example, Edwards (1963) makes the useful point that even with a four-fold increase in the risk of a common disorder the increase in concordance in monozygotic twins may be discernible only with rather large numbers.

Interracial or interethnic comparisons are a hazardous basis for conclusions about genetic factors in a common disease because one can never be certain that the groups compared are identical with reference to all pertinent exogenous factors.

Whenever a variant is shown to be significant to the pathogenesis of a common disease, the genetic factors involved in its variation can be studied by family aggregation, by twin study, and, for what they are worth, by interracial comparisons. For example, insofar as blood lipid variations are important to the pathogenesis of coronary artery disease, studies of the genetic basis of such variations are pertinent to the topic under discussion. Serum pepsinogen level was an example cited earlier by Dr. Fraser Roberts in relation to peptic ulcer. Dissecting out the *components of pathogenesis* also provides an answer to the second question in the genetics of common diseases: How do genetic factors in pathogenesis operate?

Blood-group-and-disease associations have been demonstrated in several instances, most convincingly in the case of peptic ulcer and the blood type O, nonsecretor traits. The phenomenon *association* usually has nothing to do with genetic linkage which is the location of genetic loci fairly close together on the same chromosome. Genetic linkage produces no permanent association of traits in a population. Crossing over minces up the chromosomes so that in time no association is observed. To be sure, when the linkage is close, it may be many generations before the distribution of the two traits is random. The physical basis of most *association* is, therefore, not chromosomal linkage but is rather some physiologic peculiarity, of persons of group O, for example, rendering them at greater risk to the development of peptic ulcer. Since the blood groups are simply inherited we have here one of the clearest examples of constitutional, i.e., genetic, diathesis in disease. Blood groups are, of course, not the only polymorphic traits with which association of common illness can be sought. Various serum groups such as haptoglobins, Gc, Ag and other types are further examples. Obviously if association is not demonstrated, genetic implications in the common disease are not disproved. I have put association just below genetic study of pathogenetic components because they are related matters although different in approach. The O-peptic ulcer association tells us that some feature of the type-O person is a pathogenetic component in peptic ulcer. In approach No. 4 we take a variant which, by epidemiologic or experimental means, has been shown to be a component factor in pathogenesis, and study its genetic determination. In approach No. 5 we work the other way around—take a clear-cut mendelizing trait

and test whether it has a pathogenetic component by looking for association.

If one has, in an animal other than man, a disorder seemingly identical to the common disease of man, then one may be provided with insight into possible gene-conditioned pathogenetic mechanisms in man.

Let us now consider some of the results of the application of these approaches specifically to coronary artery disease. Atherosclerosis and coronary artery disease cannot with validity be equated, yet much of the information on atherosclerosis in general is pertinent to coronary artery disease specifically and will be included.

Familial aggregation. Clinical impression (White, 1960) suggests that one should find a considerable aggregation of cases of coronary artery disease in families—more than for the garden varieties of cancer, *e.g.*, those of stomach, breast, and uterine cervix, which have been studied extensively. Unlike the cancer field, only a few family studies of coronary artery disease, each with shortcomings, have been reported. These include the studies of Gertler and White (1954), of Thomas and Cohen (1955), of Russek and Zohman (1958), and of Shanoff and colleagues (1961). Two family studies of modest size have been conducted by my colleagues Skyring *et al.* (1963) and Rose (1964). The choice of controls has been criticized in some of these studies, and problems such as better recall by proband families and biases of ascertainment are with us here as in many such studies (Schweitzer *et al.*, 1962). These criticisms aside, the results of all these studies suggest a modest familial aggregation. These are all proband studies. The population approach under way in Tecumseh was recently reported on in a preliminary manner by Epstein (1963); the familial aggregation did not look overwhelming by any means.

Twins. There are isolated reports of twins concordant (Benedict, 1958; Parade and Lehman, 1938) and discordant (Lees *et al.*, 1963) for various aspects of coronary artery disease. But these add little. The study reported yesterday by Dr. Harvald is, of course, of great interest and is, as far as I know, unique. You will recall that among 102 monozygotic twin pairs and 155 like-sex dizygotic twin pairs the concordance rate for death from coronary artery disease was essentially the same. Several questions would be of interest: What of the separate comparison of male monozygotic twins with male dizygotic twins and female monozygotic twins with female dizygotic twins? It might be of interest to look at other subgroups such as those with one twin dying of coronary artery disease under 45 years of age. And a general question comes to mind about twin studies: Does not the heritability estimate derived from twin studies vary with the environmental circumstances and is not the estimate interpretable only in a specific environmental context?

735-712 O—65——10

Many *interracial or interethnic comparisons* of the frequency of coronary artery disease have been made in recent years. In general it seems that the frequency of the disease in the groups compared becomes more nearly identical as the environmental situations become more similar (Yudkin, 1963) and it is impossible to be confident of a residual difference which might be genetic in origin.

Interesting findings were reported in New York garment workers by Epstein *et al.* (1956, 1957a, 1957b). Jewish men showed a frequency of coronary artery disease over two times greater than that in Italian men, whereas the frequency in Jewish and Italian women was about identical, at a lower value. The investigators could demonstrate little or no difference in serum cholesterol, in contemporary diet, or other variables, but one is probably still not justified (and this is Dr. Epstein's view, too) in attributing the difference to genetic difference because other very pertinent environmental variables, e.g., diet in the early decades of life, might account for the difference. I know of no interethnic data on coronary artery disease which can be confidently interpreted as supporting the genetic hypothesis.

In the area of genetic analysis directed toward pathogenetic components in coronary artery disease, studies centering around serum cholesterol and other blood lipids are the main ones. In fact, there has been an unwarranted tendency to equate the genetics of coronary artery disease with the genetics of hypercholesterolemia. I will also say something about coronary anatomy and about the coagulation mechanism.

The population approach to serum cholesterol levels would suggest that it is polygenic in its genetic basis. No bimodality is demonstrable in frequency distributions. The difficulties in recognizing two populations in distribution curves have come lately into focus with the example of blood pressure and as in that case, cholesterol shows sex and age variations which blur the picture or at least must be dealt with in the analyses. We are told by our statistical colleagues that the two populations must have certain characteristics as to separation of the means, standard deviations, and relative size to permit identification of bimodality (Murphy, 1964). There seems no doubt that there are some families in which a single major gene is responsible for hyperlipidemia, of which there are probably several genetic varieties.

But the question—polygenic or monogenic?—is perhaps irrelevant to the question which must come first: Are genetic factors of significance in determining cholesterol level? Intrafamily correlations (Shaeffer *et al.*, 1958) suggest they are significant and the studies of Osborne and Adlersberg (1958) and Osborne and De George (1959) of twins, monozygotic and dizygotic, living together and living apart, and other twin studies point more specifically in the same direction.

Coronary anatomy in terms of the pattern of gross arborization and of intercoronary anastomoses might make a good deal of difference in the localization and development of coronary occlusive disease and in the outcome of occlusive disease once developed. The findings of Schlesinger and colleagues (1940; 1949) on the relation between myocardial infarction and the pattern of coronary arborization seemed to indicate a relation, but in more recent work (Blumgart *et al.*, 1963) a relation has not been suggested. Is the pattern genetically determined to a significant extent? There are suggestions of genetic determinants in other vascular patterns: the pattern of branching of the aortic arch is different in whites and Negroes (De Garis *et al.*, 1933; McDonald and Anson, 1940; Williams and Edmonds, 1935). In monozygotic twins the arm veins have an identical pattern in 75 percent of cases whereas in dizygotic twins the figure is only about 25 percent, according to the findings of Kadanoff (1939), but the number of twin pairs studies was only 30, 8 of the 30 being monozygotic. The superficial venous pattern on the front of the thorax seems to be genetically determined to an important degree, according to Spuhler (1950). On the other hand, Platt and Lawton (1956) were not convinced, from twin studies, of a genetic factor in determination of retinal vessel pattern. The fairly extensive literature on genetic control of vascular patterns in other mammals will only be mentioned (Sawin and Nace, 1948; Sawin and Edmonds, 1949). As to the coronary pattern of man specifically, racial comparisons are about all we have to go on. Brink (1951) thought that differences in coronary arborization occurred among three racial groups living in South Africa, but Singer (1959) in a larger study in the same groups, could demonstrate no difference. Thirty years earlier Adachi (1928) could find no difference among four races.

A polemic has existed as to whether intercoronary anastomoses are an inborn endowment of the individual or whether they arise in response to demand, *i.e.*, ischemia (Blumgart, 1959, Blumgart and Zoll, 1959; Pitt, 1959). In Bantu children, Laurie and Woods (1958) found much more abundant intercoronary anastomoses than in other ethnic groups with a high frequency of coronary artery disease. The findings have been interpreted by others not as the reflection of a genetic difference but rather as a response to anemia in the Bantu child (Blumgart, 1963). Thus, in man we have no firm evidence on this matter. The extent of coronary anastomoses varies among species of animals and the genetics thereof might be studied, especially in a species which shows wide individual variability. The genetics of the gross pattern of arborization might be studied in a species with a pattern in general resembling that of man, e.g., perhaps the pig (Vastesaeger *et al.*, 1957).

The relation of hypertension and of obesity to coronary artery disease seems clear, but the genetic determination of these is a subject

I would rather not embark on now. The relation of diabetes mellitus to coronary artery disease is perhaps less clearly delineated.

The coagulation machinery (McDonald and Edgill, 1957) might be implicated in coronary artery disease in several ways. Two of these are 1) increased propensity to thrombosis and 2) if one subscribes to the view that platelet incrustation is a factor in atherogenesis, then some property of the platelet which increases its proneness to deposition. At several points the clotting mechanism and lipid metabolism seem to be interrelated although the details still require elucidation. Dr. Fraser Mustard and his colleagues in Toronto are among those who espouse an encrustation theory of atherogenesis. Murphy and Mustard (1962) studied various aspects of blood coagulation and platelet economy in subjects with atherosclerotic complications plus a family history of same and compared them with subjects free of both clinical manifestations of atherosclerosis and a positive family history. The platelet adhesive index and the platelet turnover rate showed significant differences in the two groups.

In four studies blood groups have been determined in cases of coronary artery disease and an attempt made to relate them to occurrence or clinical course. Only one study showed a significant association. This was the study of Bronte-Stewart and colleagues (1962) and here an association was found (between non-O and coronary artery disease) only in the racial group with a low prevalence of coronary artery disease. The authors suggested that in the other ethnic groups, in their study as well as in others, an association was swamped by strong environmental factors.

The study of animal analogs is illustrated in the case of atherosclerosis and coronary artery disease by two examples in birds. Pritchard and his colleagues at Bowman-Gray (1962) are working with pigeons, following up on the observation that the White Carneau has much more spontaneous atherosclerosis, including coronary occlusion, than does the Show Racer. Cross-bred pigeons have an intermediate level of susceptibility. Factors such as blood lipid difference, activity, etc., are under study, but at present the mechanism by which the genetic difference in susceptibility occurs is not known. In the Bowman-Gray pigeons, atherosclerosis was induced in both breeds by a high fat diet, but there was more fibroblastic reaction in the White Carneau, the susceptible breed. Difference in the extent of diet-induced atherosclerosis was noted by Opdyke and Ott (1954) in comparisons of White Leghorn and New Hampshire cockerel chickens. Other investigators of experimentally induced atherosclerosis have observed individual differences in the ease of induction of lesions, e.g., Kendall has noted this in the dog and Mustard in the pig.

It is evident that the study of genetic factors in a common disorder such as coronary artery disease is very difficult. The methods have not kept up with those in other areas of human genetics. We now

have methods for defining the chromosomal basis of some congenital malformations, methods for analyzing the chemical nature of genotype and phenotype all the way from the genetic code to the level of clinical manifestation, methods (albeit laborious) for mapping the chromosomes of man by linkage studies, methods for describing the dynamics of specific genes in populations. It is perhaps little wonder that one can point to a number of workers in human genetics who earlier studied disorders such as peptic ulcer, the psychoses, leukemia, diabetes, hypercholesterolemia, mental deficiency, and so on, but who have turned their attention in the last decade to these other areas in which methodology has advanced so much more rapidly. It is well to have conferences such as this to give combined thought to the role of genetic factors in common disease and the methods for their study. In time perhaps the methods in this area will catch up.

REFERENCES

1. Adachi, B., 1928. Das Arteriensystem der Japaner. Tokyo: Kenyusha Press. Cited by Singer, R., 1959.
2. Benedict, R. B., 1958. Coronary heart disease in identical female twins. Am. J. Med. *24:* 814–819.
3. Blumgart, H. L., 1959. Anatomy and functional importance of intercoronary arterial anastamoses. Circulation *20:* 812–815.
4. Blumgart, H. L., 1963. Cited in James and Keyes, 1963.
5. Blumgart, H. L., Pitt, B., Zoll, P. M., and Freiman, D. G., 1963. Anatomic factors influencing the location of coronary occlusions and development of collateral coronary circulation. *In* James and Keyes, 1963, pp. 327–339.
6. Blumgart, H. L. and Zoll, P. M., 1959. Anastamosis in the coronary circulation. Lancet *1:* 207.
7. Brink, A. J., 1951. South African Jour. Clin. Sci. *2:* 288. Cited by Singer, 1959.
8. Bronte-Stewart, B., Botha, M. C., and Krut, L. H., 1962. ABO blood groups in relation to ischaemic heart disease. Brit. Med. J. *1:* 1646–1650.
9. Davenport, C. B., 1916. The hereditary factor in pellagra. Arch. Int. Med. *18:* 4–31.
10. De Garis, C. F., Black, I. H., and Riemenschneider, E. A., 1933. Patterns of the aortic arch in American white and Negro stocks, with comparative notes on certain other mammals. J. Anat. *67:* 599–619.
11. Edwards, J. H., 1960. The simulation of Mendelism. Acta Genet. et Statis. Med. *10:* 63–70.
12. Edwards, J. H., 1963. The genetic basis of common disease. Am. J. Med. *34:* 627–638.
13. Epstein, F. H., 1963. Hereditary aspects of coronary artery disease. *In* Symposium on "Genetics and Heart Disease," Philadelphia, Jan. 24–26.
14. Epstein, F. H., Simpson, R., and Boas, E. P., 1956. Relations between diet and atherosclerosis among a working population of different ethnic origin. Am. J. Clin. Nutrition *4:* 10–19.
15. Epstein, F. H., Boas, E. P., and Simpson, R., 1957a. The epidemiology of atherosclerosis among a random sample of clothing workers of different ethnic origins in New York City. I. Prevalence of atherosclerosis and some associated characteristics. J. Chronic Dis. *5:* 300–328.

16. Epstein, F. H., Simpson, R., and Boas, E. P., 1957b. The epidemiology of atherosclerosis among a random sample of clothing workers of different ethnic origins in New York City. II. Associations between manifest atherosclerosis, serum lipid levels, blood pressure, overweight, and some other variables. J. Chronic Dis. *5:* 329–341.
17. Gertler, M. M. and White, P. D., 1954. Coronary Heart Disease in Young Adults: A Multidisciplinary Study. Cambridge: Harvard University Press.
18. Haenzel, W., 1959. Some problems in the estimation of familial risks of disease. J. Nat. Cancer Inst. *23:* 487–505.
19. Horn, D., Courts, F. A., Taylor, R. M., and Solomon, E. S., 1959. Cigarette smoking among high school students. Am. J. Public Health *49:* 1497–1511.
20. Jaines, T. N. and Keyes, J. W., editors, 1963. The Etiology of Myocardial Infarction. Boston: Little, Brown and Co.
21. Kadanoff, D., 1939. Untersuchungen über die Unterschiede in der Verästelung der Hautnerven und -venen der oberen Extremität bei ein- und zweieiigen Zwillingen. Ztschr. f. Morphol. u. Anthropol. *38:* 73–89.
22. Kendall, F. E., New York: Personal communication.
23. Laurie, W. and Woods, J. D., 1958. Anastamosis in the coronary circulation. Lancet *2:* 812–816.
24. Lees, R. S., Canellos, G. P., Rosenberg, I. H., and Hatch, F. T., 1963. Myocardial infarction in one of a pair of 27-year-old identical male twins. Am. J. Med. *34:* 741–746.
25. Lilienfeld, A. M., 1959. A methodological problem in testing a recessive genetic hypothesis in human disease. Am. J. Publ. Health *49:* 199–204.
26. McDonald, J. J. and Anson, B. J., 1940. Variations in the origin of arteries derived from the aortic arch in American whites and Negroes. Am. J. Phys. Anthrop. *27:* 91–107.
27. McDonald, L. and Edgill, M., 1957. Coagulability of the blood in ischaemic heart disease. Lancet *2:* 457–460.
28. Muncey, E. B., 1916. A study of the heredity of pellagra in Spartanburg County, South Carolina. Arch. Int. Med. *18:* 32–75.
29. Murphy, E. A., 1964. One cause? Many causes? The argument from the bimodal distribution. Fallacies and pitfalls. *J. Chron. Dis. 17:* 301–324.
30. Murphy, E. A. and Mustard, J. F., 1962. Coagulation tests and platelet economy in atherosclerotic and control subjects. Circulation *25:* 114–125.
31. Mustard, J. F., Toronto: Personal communication.
32. Opdyke, D. F. and Ott, W. H., 1954. Influence of source of cholesterol, grade of cotton-seed oil, and breed on experimental avian atherosclerosis. Proc. Soc. Exp. Biol. and Med. *85:* 414–415.
33. Osborne, R. H. and Adlersberg, D., 1958. Serum lipids in adult twins. Science *127:* 1204.
34. Osborne, R. H. and DeGeorge, F. V., 1959. Genetic Basis of Morphological Variation. An Evaluation and Application of the Twin Study Method. Cambridge: Harvard University Press.
35. Parade, G. W. and Lehman, W., 1938. Angina pectoris bei erbgleichen Zwillingen. Klin. Wchnschr. *17:* 1036–1040.
36. Penrose, L. S., 1953. The genetical basis of common disease. Acta Genet. et Statis. Med. *4:* 257–265.
37. Pitt, B., 1959. Interarterial coronary anastomoses: occurrence in normal hearts and in certain pathologic conditions. Circulation *20:* 816–822.
38. Platt, R. and Lawton, R., 1956. The retinae of monovular and binovular twins. Ann. Human Genetics *21:* 132–134.

39. Pritchard, R. W., Clarkson, T. B., Lofland, H. B., Goodman, H. O., Herndon, C. N., and Netsky, M. G., 1962. Studies on the atherosclerotic pigeon. J. Amer. Med. Assn. *179:* 49–52.
40. Rose, G., 1964. Familial patterns in ischaemic heart disease. Brit. J. Prev. Soc. Med. *18:* 75–80.
41. Russek, H. I. and Zohman, B. L., 1958. Relative significance of heredity, diet and occupational stress in coronary heart disease of young adults. Amer. J. Med. Sci. *235:* 266–277.
42. Sawin, P. B. and Edmonds, H. W., 1949. Aortic arch variations in relation to regionally specific growth differences. Anat. Record *105:* 377–397.
43. Sawin, P. B. and Nace, M. A. G. 1948. Morphogenetic studies of the rabbit. V. Inheritance of an asymmetrical vascular pattern. J. Morphol. *82:* 331–354.
44. Schlesinger, M. J., 1940. Relation of anatomic pattern to pathologic conditions of the coronary arteries. Arch. Path. *30:* 403–415.
45. Schlesinger, M. J., Zoll, P. M., and Wessler, S., 1949. The conus artery: a third coronary artery. Am. Heart J. *38:* 823–836.
46. Schweitzer, M. D., Clark, E. G., Gearing, F. R., and Perera, G. A., 1962. Genetic factors in primary hypertension and coronary artery disease—a reappraisal. J. Chron. Dis. *15:* 1093–1108.
47. Shaeffer, L. E., Adlersberg, D., and Steinberg, A. G., 1958. Heredity, environment and serum cholesterol. A study of 201 healthy families. Circulation *17:* 537–542.
48. Shanoff, H. M., Little, A., Murphy, E. A., and Rykert, H. E., 1961. Studies of male survivors of myocardial infarction due to "essential" atherosclerosis. Canad. Med. Assn. J. *84:* 519–530.
49. Singer, R., 1959. The coronary arteries of the Bantu heart. South African Med. J. *33:* 310–315.
50. Skyring, A., Modan, B., Crocetti, A., and Hammerstrom, C., 1963. Some epidemiological and genetic aspects of coronary heart disease. Report of a pilot study. J. Chronic Dis. *16*: 1267–1279.
51. Spuhler, J. N., 1950. Genetics of three normal morphological variations: pattern of superficial veins of the anterior thorax, peroneus tertius muscle and number of vallate papillae. Cold Spring Harbor Symp. Quant. Biol. *15:* 175–189.
52. Thomas, C. B. and Cohen, B. H., 1955. The familial occurrence of hypertension and coronary artery disease, with observations concerning obesity and diabetes. Ann. Int. Med. *42:* 90–127.
53. Vastesaeger, M. M., Van der Straeten, P. P., Friart, J., Candaele, G., Ghys, A., and Bernard, R. M., 1957. Les anastomoses intercoronariennes telle qu'elles apparaissent à la coronarographie post mortem. Acta Cardiologica *12:* 365–401.
54. White, P. D., 1960. The importance of heredity in coronary heart disease. Circulation *22:* 296–300.
55. Williams, G. D. and Edmonds, H. W., 1935. Variations in the arrangement of the branches arising from the aortic arch in American whites and Negroes (a second study). Anat. Rec. *62:* 139–146.
56. Yudkin, J. 1963. Racial and ethnic factors in the etiology of myocardial infarction. *In* James and Keyes, 1963, pp. 3–11.

The Geographic and Genetic Characteristics of Amyotrophic Lateral Sclerosis and Other Selected Chronic Neurological Diseases

By Leonard T. Kurland,[1] M.D., Dr. P.H.,
Chief, Epidemiology Branch, National Institute of Neurological Diseases and Blindness, National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare, Bethesda, Md.

INTRODUCTION

For most chronic diseases of the nervous system we are unable to say whether the principal cause is exogenous or genetic. Since the direction taken in research is often influenced by the prevailing etiologic viewpoint, systematic efforts are being made to assemble a fund of knowledge through descriptive epidemiology and genetic procedures to guide us, for each disorder, to the most appropriate position in the heredity-environment "continuum."

For many of the neurologic disorders the first step in this process has been to determine the pattern of geographic distribution. The pattern, if uneven or one with focalization of cases, generally conjures exciting prospects for the epidemiologist; if one of random distribution on a national level or within a community, it may pose a frustrating challenge until some breakthrough occurs at the clinical or laboratory level.

For a number of disorders which can be readily diagnosed, and have a high case fatality rate, and for which there are specific rubrics, age-adjusted mortality rates have been compared on an international basis (Goldberg and Kurland, 1962). For those disorders in which the rate is high enough, comparisons have been made by State and

[1] Present address: Division of Biometry and Medical Statistics, Mayo Clinic, Rochester, Minn.

Province or by regions in the United States and Canada. Where the pattern of mortality distribution and other characteristics permit, intensive morbidity surveys and "case-control" comparisons have been conducted in communities which are of appropriate size, level of medical practice and geographic distribution, in the hope of identifying some focalization by family, race, socio-economic status, etc., within the community worthy of more specific genetic or environmental study (Kurland, 1958).

Geographic isolates in which a "new" disease is discovered or in which a well-known illness is unusually prevalent are particularly intriguing to the epidemiologist and geneticist. The elucidation of a "new" disease may improve our understanding of basic physiological processes or pathogeneses; the discovery of an "epidemic" focus of a chronic disease which is ordinarily uncommon may provide an ideal set of conditions for comparisons which can never be attained in ordinary clinical studies. In recent years in neurology, several foci which carry such exotic names as "Minamata disease," "kuru," "Jamaican neuropathy," "lytico," and "bodig" have been studied intensively and productively. All have the following features in common: a high incidence of a central nervous system (CNS) disease which is generally fatal; an isolated geographic setting; some familial aggregation of cases; and, early in the investigation, uncertainty of the role of heredity. If Minamata disease, the only disorder which appears to have been "solved," is any indication, the etiologic and pathogenic mechanisms will be as intriguing as the names themselves. After several years of study at the University of Kumamoto, Japan, a slowly progressive brain stem and encephalitic process which affected fish-eating birds and cats as well as villagers, was found to be due to an organic mercury compound present in the fish of a small bay near the city of Minamata, Japan. The compound was derived from a pollutant of a nearby industrial plant (Kurland *et al.*, 1960). Kuru, a disease which appears to be unique among the primitive Fore tribe in the remote highlands of New Guinea, accounts for a phenomenal mortality rate; there is still considerable controversy over whether the disease is primarily due to an exogenous or hereditary mechanism (Gajdusek and Zigas, 1959; Fischer, 1961; Williams *et al.*, 1964). Jamaican neuropathy is a progressive disease characterized by diffuse involvement of the central nervous system with some selection for the cortico-spinal and other tracts and the optic nerve (Cruickshank *et al.*, 1961). Unfortunately, little is known about the geographic distribution on Jamaica and throughout the islands of the Caribbean, nor are other epidemiologic and genetic features described. The foci of lytico and bodig will be considered below for they may be a key to the understanding of the important problem of amyotropic lateral sclerosis.

General Features

Amyotrophic lateral sclerosis, which for convenience will be referred to as ALS, is a progressive fatal disorder of adults due to degeneration of neurons in several areas of the CNS; in the spinal cord, resulting in the clinical picture of progressive muscular atrophy; in the bulbar motor nuclei, resulting in the picture of progressive bulbar palsy; and in the motor cortex of the cerebrum, resulting in the upper motor neuron components of pseudobulbar palsy and lateral sclerosis which are usually present (Friedman and Freedman, 1950). Clinically, the disorder in the indigenous Chamorro people of Guam, to which we turn shortly, is indistinguishable from that seen in other lands and is characterized by progressive weakness and wasting of the bulbar and skeletal musculature, often with hyperreflexia and pathological reflexes. There are no signs of sensory, sphincter, cerebellar or mental dysfunction in the "classical" case. As the thoracic and bulbar musculature and other muscles become affected, respiratory weakness may lead to pneumonia or pulmonary decompensation; bulbar paralysis may result in dysphagia and progressive inanition, or it may be responsible for aspiration of food and asphyxia, or broncial pneumonia.

In the United States there is no appreciable difference in the mortality rate by regions of the country. The rate in the northeast, southeast, northwest and southwest, and in the state of Hawaii is approximately 0.7 per 100,000 population annually. These rates are derived from those certificates in which ALS or motor neuron disease is listed as the primary cause of death. The age-adjusted rates for Canada, several of the Scandinavian countries, Holland, Great Britain, Australia and Japan also show that the mortality rates are remarkably alike as the age-adjusted rates for these countries all fall within a narrow range between 0.7 and 1.0 per 100,000 population. The sex ratios for these same countries have also been remarkably consistent: approximately 1.6 males per female (Kurland, 1957).

The death ratio is another way of representing the risk from this cause, *i.e.*, the number of deaths from ALS compared to the total deaths. For example, the death ratio for ALS in the adult population of the United States is approximately 1 per thousand; *i.e.*, 1 out of each 1,000 deaths is due to ALS. However, among the 40,000 Chamorros of Guam and the other Mariana Islands, ALS is responsible for about 10 percent of the deaths in adults. This phenomenal concentration of disease was observed by several physicians in the U.S. Navy during and after World War II and resulted in a series of investigations aimed at explaining this focus and, if possible, the nature of ALS itself (Arnold *et al.*, 1953; Koerner, 1952; and Kurland and Mulder, 1954). The results and conclusions of the earlier studies will first be reviewed; the effect of recent observations which have

forced a reconsideration of some of those earlier conclusions will then be discussed and should serve as an illustration of the need for thoroughness in population studies and caution in formulating etiologic hypotheses.

Early Studies in the Mariana Islands

Guam, the largest and southernmost of the Mariana Islands, is a territory of the United States inhabited by the Chamorros and a few thousand other Western Pacific peoples and about 35,000 military personnel, dependents, and other statesiders. The disease among the Chamorros was described in the initial studies as a "classical" form. The earlier pathologic studies demonstrated the typical degenerative findings in the upper and lower motor neurons (Sayre, 1957). On Guam more than 100 patients were identified during the early surveys; all such cases were among the Chamorros. In the preliminary surveys of the neighboring islands, 5 cases were noted among the 5,000 Chamorros on Rota, Tinian and Saipan and 1 case was observed among the 1,000 Carolinians on Saipan. The mother of this latter patient had migrated from the Caroline Islands to Saipan via Guam about 50 years earlier and the patient was born on Saipan. It was said that the patient's father was also a Carolinian.

Thus the disease at first appeared to be limited to persons who had at least one Chamorro parent, with the exception of the one Carolinian case on Saipan and even here there was some question as to paternity. In those few instances on Guam where there was mixed parentage, the Chamorro parent was the mother and the father was Filipino, Japanese or Caucasian. At the time of these surveys we tended to discount the one Carolinian case and since there were few marriages between Chamorro men and non-Chamorro women, we were not concerned by the fact that in mixed marriages with affected offspring on Guam, the mother was invariably Chamorro.

Surveys in California disclosed that among several hundred Chamorros who had migrated to California after World War II the frequency of ALS was the same as among their Chamorro brethren on Guam (*i.e.*, about 1 percent prevalence ratio and about 10 percent mortality ratio). A few cases before World War II were disclosed and these were in men who had been in California for over 20 years before they were affected (Torres *et al.*, 1957).

Although familial aggregation of cases was recognized on Guam it proved difficult to obtain reliable pedigrees (Koerner, 1952; Mulder and Kurland, 1954) and no simple consistent genetic pattern could be identified. Most civil and church records had been destroyed during World War II; the older population was largely illiterate and had maintained few records of a family nature. Illegitimacy, at least as reported on civil records, was appreciable. Even though

the disease was common enough to be known as "lytico" (from the Spanish "paralytico") there was frequently denial of the disease by relatives even when the family had accurate information on such cases. This is due to the stigma of ALS, since according to Chamorro folklore, ALS specifically is retribution for sins which the individual or his ancestors committed against the Church (Kurland and Mulder, 1954; Kurland, 1957). In the village of Umatac, where the incidence rate of ALS is the highest on the island, it was possible to trace many current cases of ALS back through several generations to the ancestor who allegedly brought the curse upon the people of the island by his repeated petty larcenies of church property (Mulder and Kurland, 1954).

During the early stages of the program on Guam a concomitant evaluation of ALS at the Mayo Clinic disclosed that about 1 in 15 cases gave a positive family history. Furthermore, when such families were intensively studied, pedigrees could be obtained for several generations in which a pattern compatible with dominant inheritance and a high degree of penetrance was observed (Kurland and Mulder, 1955).

A review of the literature revealed that even in an early description of ALS by Aran in 1850, 1 of the original 8 cases had an extensive family history. The review of the literature and the study at the Mayo Clinic in 1954 provided 26 pedigrees, all of which are compatible with the dominant pattern and a high degree of penetrance. One of the most striking pedigrees is that by Osler, brought up to date by Brown, which demonstrates 17 cases, including one set of twins, (zygosity undetermined) who were affected over 4 generations (Brown, 1951). An additional dozen or more families have been reported in the past decade, almost all showing the same pattern (Hasaerts-Van Geertruyden, 1955; Boudin and Barbizet, 1956; Gervas, 1959; Haberlandt and Glanville, 1962; Bunina and Kanowalow, 1962; and Sercl and Kovarik, 1963).

The high family incidence is not limited to a brief period of time when various members might be exposed to a common agent, but is spread over generations. In some cases the disease continued to occur in families in which some members had migrated from Europe to America.

An analysis of the collected pedigrees reveals that the range of ages of onset and the duration of illness were similar to those of the sporadic cases in the United States and the cases on Guam also (table 1). However, in contrast to the observed preponderance of males in the other groups, the familial cases have shown a 1:1 sex ratio and the mean age of onset of the Guam cases is apparently less than the sporadic cases.

It was noted by Kurland and Mulder (1955) that in the sibships with at least one affected member, about 50 percent of the siblings

TABLE 1.—*Selected clinical and pathological features of amyotrophic lateral sclerosis*

	"Sporadic" cases	"Familial" cases	"Mariana" cases
Approximate mean age onset	53 yrs	47 yrs*	45 yrs**
(usual range)	(25–65)	(25–65)	(25–65)
Sex ratio M:F	1.6:1	1:1	2:1**
Approximate mean duration, years (usual range) †	3.0 (1–8)	3.0 (1–10)	3.0 (1–8)
Clinical variants:			
Extrapyramidal	Rare	Occasional	Frequent (if cases with both ALS & PD are included)
"Dementia"	Rare	Occasional	"
Sensory changes	Not detected	Not detected	Not detected
Histochemical changes, skin (collagen)	Over 50 percent	Infrequent(?)	Over 50 percent
Pathology:			
(1) Posterior column demyelination and column of Clarke changes.	Rare	50 percent (estimated)	Rare
(2) Neurofibrillary changes before age 60 years.	Rare	Rare	Common
(3) Granulovacuolar bodies	Rare	Rare	Occasional

*Intrafamilial variation is less than interfamilial variation.
**Cases of P–D have a mean age of about 50 years and a sex ratio of about 3:1.
†Since 1950, the reported mean duration has increased about 1 year in the Mariana cases to 4.0 years; this may apply to other groups as well and may be related to improved supportive therapy or the earlier recognition and more accurate establishment of the date of onset.

(88 affected; 91 unaffected) had developed the disorder. When they allowed for the possible bias of ascertainment of sibships through affected individuals the ratio of affected to unaffected was somewhat less than 50 percent. Only about 1 in 15 patients had reported another case in the family and many of the histories were negative after careful interview and documentation. On the other hand, the striking collection of pedigrees which were reported; and which might have led to the conclusion that all cases of ALS were genetic; could also be misleading, for there is a tendency among physicians to report or recall not the usual kindred but the exceptional ones.

At the time of the ALS investigations on Guam in 1956 and 1957 we were unable to associate any environmental factor with the disorder, although the disease incidence seemed to vary inversely with the degree of urbanization. We concluded from all the data collected that ALS was presumably inherited as an autosomal dominant trait with a high penetrance, at least in the Chamorros. Our reasoning at that time included the following considerations: ALS was exceedingly prevalent in the Chamorros whether they lived on Guam, the other Mariana Islands or in California. With the exception of one case, the disease was not known in the other ethnic groups on the same islands. Although the pedigrees could not be analyzed satisfactorily, familial aggregations often over several generations were known. Numerous pedigrees of ALS had been reported elsewhere in which the disease behaved as an autosomal dominant trait. Furthermore, the disease was particularly selective for one "system" in the CNS. If ALS were due to a dominant trait on Guam the gene frequency would have to be exceedingly high. This would be difficult to explain for there was no obvious selective advantage for ALS. Even though the disease had its onset in adulthood, half the cases developed during

the reproductive years. It was suggested that in order to attain the high gene frequency a genetic hypothesis would require, the gene responsible for the disease should have been introduced at a time when the population was relatively small and, through the effect of drift or some selective survival mechanism, could then have spread to the extent that ALS was recognized as a familial disorder even in the early 19th century. We concluded that either the disease was introduced early in the human history of Guam, perhaps 2,000 years ago, or after the Spanish colonization which began about 1667. There is a suggestion from the writings of the Jesuits who arrived with the first Spanish colonists that the population was healthy and long-lived so it seems less likely that ALS was prevalent at that time. During the period from 1670 to 1700 the introduction of infectious diseases into the population, war with the colonists; and famine almost wiped out the Chamorros. It has been suggested that ALS was inadvertently introduced into the small residual population during the period of the colonization and that the repeated epidemics of smallpox from 1700 to 1860 might have resulted in selective survival of those who had inherited the tendency for ALS. Such selective survival might be present if ALS had been derived from one or more Spaniards or other Europeans, and fortuitiously was associated with a higher natural immunity to the smallpox which periodically decimated the population (Kurland and Mulder, 1954; Kurland, 1957).

More Recent Studies

Since 1957, there have been several developments which have influenced our thinking about the possibility of a simple genetic basis for the disease on Guam. These are:

(1) the realization that another neurological disorder formerly considered as post-encephalitic Parkinsonism and referred to as Parkinsonism-dementia complex (PD); is also highly prevalent among the Chamorros (Kurland *et al.*, 1961) and probably is part of the ALS spectrum on Guam;

(2) the study which revealed changes in the collagen structure of the skin in patients with ALS and PD (Fullmer *et al.*, 1960);

(3) the observations that diaphysial aclasis (multiple exostosis), gout and diabetes were highly prevalent in the population (Krooth *et al.*, 1961; O'Brien *et al.*, 1964);

(4) on the island of Saipan the observation of additional cases of ALS among individuals whose mothers were Carolinian and whose fathers were presumably Carolinian (Mathai, 1963);

(5) a new focus of ALS in another area of the western Pacific—this one on the Kii Peninsula of Japan (Kimura *et al.*, 1961);

(6) a possible new focus of ALS in southwestern New Guinea (Gajdusek, 1963).

Parkinsonism-dementia (PD), known on Guam as "raput" or "bodig," has, as its main features, an organic type of mental deterioration, a reduced spontaneous motor activity referred to as akinesia, and hyperreflexia (Hirano *et al.*, 1961). This is due to the neural atrophy of the cerebral cortex, particularly the fronto-temporal regions, and to degeneration of the basal ganglia, particularly the pigmented areas such as the substantia nigra (Hirano *et al.*, 1961). This disease accounts for about 7 percent of adult mortality on the island and has a mean age at onset of about 50 years, with a duration of about 4 years from onset to death. In 1 out of 6 cases there is amyotrophy due to anterior horn cell degeneration, the characteristic feature of ALS. Careful study has shown that advanced ALS cases occasionally develop evidence of the PD features. This association of motor neuron, extrapyramidal, and dementia features is not unique to the "Guam" cases, for there are a few reports of individual and familial cases of ALS in which Parkinsonism, dementia or both are present (Bonduelle *et al.*, 1959; Légrand *et al.*, 1959; Hanau, 1960; Patrikios, 1951; Van Bogaert and Radermecker, 1954; and Robertson, 1953).

In PD, and to a lesser extent in the ALS of Guam, widespread neurofibrillary changes occur but are especially severe in Sommer's sector of the hippocampus. In addition, in those areas in which fibrillary changes in the neurones are severe, granulovacuolar bodies or a type of intracytoplasmic "inclusion" body appear; their exact pathogenesis with respect to a degenerative or an infectious process is still unknown; although inclusion bodies are equated with a virus by some workers (Konowalow *et al.*, 1963; Zilber *et al.*, 1963).

As we became more adept in approaching the people in the islands it was possible to obtain family histories in which siblings or parents of patients were known to be affected by ALS or by PD. Although most Chamorros differentiate ALS from PD clinically and do not consider PD as having the same degree of stigma as ALS, a close relationship between the two conditions is suggested by the similarities of the age distribution, sex ratio, duration, some pathological and clinical features and the frequent history of both disorders in the same family. There is also a suggestion that the incidence rates of the "two" disorders fluctuate in a parallel fashion in some of the villages on Guam and Rota. It now seems highly likely that ALS and PD among the Chamorros are clinical variants of one disease mechanism; the factors responsible for the spectrum of clinical manifestations are unknown. The unusual opportunity to observe the population of the Marianas in an almost complete medical and pathological sense perhaps provides us with a total view of the spectrum of a disease which may not be available to most clinicians or pathologists elsewhere.

The dermal changes which are present in about 60 percent of the ALS cases both in the continental United States and Guam and in the

PD cases include increased mucopolysaccharide, connective tissue disorientation, elastosis, and disorganization of the collagen structure (Fullmer *et al.*, 1960). The basis for these changes is uncertain but they need not be associated with severe muscle wasting under the site of the skin biopsy.

The prevalence rate of diaphysial aclasis is about 1 per 1,000 on Guam but usually occurs in a familial pattern compatible with dominant inheritance (Krooth *et al.*, 1961). Evidence has mounted that hyperglycemia and hyperuricemia are also unduly prevalent in the Chamorros and the Carolinians of the Marianas. Although the association of these disorders with the neurological disease is uncertain, the possibility of a common toxic factor must be considered.

In the last 4 years at least 3 additional cases of ALS were diagnosed among the Carolinians on Saipan and 4 other suspected cases have occurred in non-Chamorros, including a civilian statesider and 1 Filipino who had been living in the native villages for over 10 years. In spite of the fact that this number does not necessarily exceed the expected number for the entire non-Chamorro population, we have been concerned because the non-military stateside and Filipino populations are small and two of these patients also showed some clinical features of PD. Although the possibility remains that the parentage of the Carolinian cases may be partly Chamorro, as the number has increased year after year, this explanation has become less tenable and has forced us to reconsider the possibility that an exogenous etiologic factor plays more of a role than we had previously assumed. This exogenous condition might be exposure to something which may have been a part of the Chamorro culture and which was adopted in recent decades by the Carolinian immigrants to Saipan. However, before this can be established, it may be necessary to conduct intensive surveys in the "pure" Carolinian population on the islands south of Guam to determine whether the disease is also unusually prevalent in these islands.

The ALS cases in the Kii Peninsula of Japan cannot be distinguished clinically from those on Guam. Only one has come to autopsy and this patient had the same neurofibrillary changes observed in the Guamanian cases (Shiraki, 1962). There is a possibility that the ALS in the remote villages of the Kii Peninsula is related to a common origin of the Chamorros and these Japanese, since it is conceivable that Chamorros swept out to sea by the currents in the vicinity of Guam and Saipan could have been carried to this part of Japan in some remote period. However, no genetic association has been established between these peoples, and it seems more likely that there is some environmental factor common to them.

In one other focus, on the south coast of western New Guinea, Gajdusek recently observed a high incidence of muscular atrophy of the hands of adolescents and young adults, and the presence of a

735-712 O—65——11

suspiciously high incidence of a motor neuron disease. The latter disorder is probably ALS but requires pathological confirmation.

These developments have caused us to intensify our studies of the clinical and pathological features of ALS to determine whether we are dealing with a single entity or several diseases. For the moment I shall refer to three forms of the disease; namely, the "sporadic" cases, such as those in the United States in which there is no familial history; the "familial" cases as they have been reported in the United States and from many other countries; and the "Mariana Islands" cases, particularly those of the Chamorros of Guam.

In table 1 a number of features of these three groups are compared. Although the range of the age of onset is similar, there is the tendency, according to Reed *et al.*, for the "Mariana" cases to occur at a somewhat younger age than the "sporadic" cases reported in the Mayo Clinic series by Mulder (1957) and in the analysis of the United States mortality data by Van Nuis (1964). The duration is similar but the sex ratio for the Mariana and sporadic cases differs from that of the 1 to 1 ratio observed in familial cases. The PD features which have been described in the Guam cases are noted infrequently in the other two groups. Clinically, none of the groups show any sensory changes due to involvement of the posterior columns but about half of the familial cases which have come to my attention show posterior column demyelination and column of Clarke cell loss. The posterior column changes resemble those seen in the hereditary ataxias. This suggests that in the "spectrum" of neurological system diseases some of the familial ALS cases may shade into the ataxia group. If the Guam cases also prove to be primarily genetic, this form of ALS would presumably blend into the Parkinsonism and presenile dementia groups of a neurologic disease spectrum.

The histochemical changes observed in the skin of Mariana and sporadic cases has not been noted in the few cases of familial ALS which have been studied (Fullmer *et al.*, 1960).

The neurofibrillary changes and granulovacuolar bodies which are seen so frequently in the Guam cases are rarely observed in either the familial or the sporadic cases and, in the latter, only rarely before the age of 60 years. In contrast to this, the Guam cases show such changes even in the youngest of the ALS cases observed at autopsy (Hirano and Zimmerman, 1962).

It remains to be seen whether I have been unduly emphasizing differences in the three groups instead of the many features of a clinical and pathological nature which they share. The question remains as to whether the disease on Guam and the other Pacific Islands is fundamentally different from that seen in the Western populations. There is also a need to determine whether we are observing, perhaps still incompletely, the spectrum of a single disease with its variants in this phenomenal focus in the Marianas. If it is

a single disease, is it so prevalent here because of an increased exposure to some subtle but highly concentrated toxic, infectious or other exogenous agent which is present in some of these islands? Or is it due to a common genetic trait in these apparently different ethnic groups in the long chain of western Pacific Islands, or a combination of increased genetic susceptibility and a prevalent exogenous factor?

Obviously, every effort should be made to clarify the situation on Guam and Saipan where conditions are well suited for intensive epidemiologic studies. This will include an increased effort to conduct total village surveys on Guam, including non-Chamorro residents, to determine precisely the pattern of distribution of ALS and PD and other disorders. Comparisons between villages of high and low incidence and between cases and an appropriate comparison group must also be undertaken. On a broader scale, the pattern of distribution throughout the western Pacific Islands must be clarified and exogenous factors common to ethnically different but affected people must be brought to light. Also, features of islands with differing incidence rates must be studied—particularly those inhabited by people who appear to be ethnically similar.

The search for a possible toxic agent continues. At the moment we are trying to determine whether there may be an agent which has the capacity to act as do the aminonitrile compounds in human and experimental lathyrism where such compounds can, in the appropriate animal species and at the appropriate age, produce defects including neuronal degeneration, collagen changes, or a form of multiple exostosis, even though the latter differs in site from the exostosis described by Krooth *et al.* on Guam (1961). Whether the high incidence of gout and diabetes is related to the neurological disorders and whether a toxic basis for these exists also requires elucidation.

A search has not yet disclosed any genetic markers which are nonrandomly associated with the neurological diseases.

Needless to say, efforts must be continued to clarify the clinical, pathological, and epidemiological factors and the relationships of the sporadic, familial and Mariana cases of ALS.

OTHER NEUROLOGICAL DISORDERS

For other illustrations of problems in the etiologic interpretations of broad geographic patterns of disease incidence, data are available for several other afflictions of the nervous system. Because their geographic distributions contrast with those of the ALS group of diseases which have been described, multiple sclerosis and primary brain tumors have been selected for this presentation.

Multiple Sclerosis

Multiple sclerosis has a unique but still unexplained geographic distribution. The earliest of several extensive international mortality

studies indicated that the rate decreases as one approaches the equator from either direction (Limburg, 1950). Later studies which have introduced some refinements have shown this same pattern (Acheson, 1961; Kurland *et al.*, 1957; Kurland and Reed, 1964).

Morbidity surveys reveal that the prevalence rate for several cities in Canada and the northern United States is several times greater than the rate in the southern cities (Kurland and Westlund, 1954). A survey in Israel by Alter *et al.* (1962) demonstrated that the prevalence of multiple sclerosis was appreciably greater among immigrants to Israel from northern European countries than among immigrants from the Mediterranean or Near East countries. Alter also observed that the rate among the Israeli-born population was low regardless of the country of origin of the parents. None of the community surveys completed in North America have revealed focalization of the disease within the community; there is no apparent selection by race, occupation or various socio-economic measurements (Kurland and Westlund, 1954). Acheson *et al.* (1960) attempted to identify meaningful variables correlated with latitude by studying the distribution of men discharged from Veterans Administration hospitals with the diagnosis of multiple sclerosis. Multiple regression analyses were made on the basis of degrees north latitude of birthplace, average hours of sunshine, average annual degree day (an index of severity of winter) and average daily December solar radiation. The highest correlations of multiple sclerosis prevalence were with average total annual hours of sunshine and with average daily December solar radiation. The fact that the correlation between multiple sclerosis and sunshine was a negative one suggests that if sunshine exerts an effect it must be preventive or protective rather than harmful.

The geographic variation of multiple sclerosis and the failure to observe any significant focalization in community studies are more suggestive of an environmental than a genetic basis but this fortunately has not discouraged a number of genetic studies.

There are serious problems, none of them unique, in genetic studies of multiple sclerosis: (1) There is no specific diagnostic test for multiple sclerosis and the error of diagnosis, particularly in the earlier stages, may be appreciable (Stazio *et al.*, 1964). (2) It is particularly difficult to differentiate multiple sclerosis on clinical grounds from some of the progressive ataxias which are generally recognized as being hereditary. (3) The prevailing view among clinicians is that multiple sclerosis is not hereditary so that a family history of a similar disorder may discourage the physician from making a diagnosis of multiple sclerosis.

The data which have been presented and which have sometimes been interpreted as suggestive of a genetic basis for multiple sclerosis are the following types: (1) The percentage of "positive family his-

tories" in a consecutive series of cases. In most of these the family history is inadequate; there is usually a failure to define the size of the "family," and there is an absence of any clearcut familial pattern (Pratt *et al.*, 1951). (2) Twin studies have shown an occasional concordance for monozygotic or dizygotic pairs (Mackay and Myrianthopoulos, 1958). In one series reviewed by Müller (1953) in which there was no apparent bias in selection of twins, 50 pairs, at least 1 of which had multiple sclerosis, were examined. All of the 14 monozygotic pairs and all but 1 pair of dizygotic twins were discordant as far as multiple sclerosis was concerned. (3) Family members within a specified degree of relationship have been enumerated and examined and the number of cases compared with an "expected" figure, based on mortality data or from population surveys (Myrianthopoulos and Mackay, 1960). Since the latter surveys are actually derived from a review of various sources of medical information and the population has not been examined to the same degree as the family members, this type of comparison is not without bias. These studies generally show that about 0.5–1.5 percent of siblings are affected and about 0.5 percent of the parents are affected as compared to an "expected" ratio of 1/2,000 in the total population. The lack of age adjustment, the problems of diagnosis on a comparison basis and the inadequacies of the available rates for the population at large could account for some of this discrepancy; the question of diagnosis is a particularly important one since the inclusion of only one or two hereditary ataxias in a small group of cases could provide a statistically significant difference between the number of cases in the family and the number "expected" (Kurland *et al.*, 1955).

Thus, the data for multiple sclerosis taken as a whole show a slight tendency to familial aggregation of cases. Whether the familial aggregation which is observed is due, at least in part, to a genetic factor or to exposure to some common agent remains to be seen. Nonetheless, future etiologic hypotheses for multiple sclerosis should be consistent with the remarkable geographic distribution of the disease.

Brain Tumors

The latest primary brain tumor mortality rates from all countries from which such data were available have revealed no appreciable geographic differences (Goldberg and Kurland, 1962). Within the United States by states there appears to be an even distribution for the brain tumor rubrics taken as a whole (Kurland *et al.*, 1962).

Aside from the tumors which accompany some of the phakomatoses, including von Recklinghausen's or von Hippel Lindau's disease, which represent developmental anomalies in which heredity plays an important role, there is no clearcut evidence for the role of a genetic mechanism in primary brain tumors. The inadequacy of most family

histories with respect to brain tumor may be an important factor in the apparent lack of familial aggregation of meningiomas and gliomas, for neurosurgery has only recently been providing the diagnostic procedures which are necessary to differentiate clearly the symptoms of brain tumor from those of other central nervous system disorders. Nevertheless, there are remarkably few case reports of such tumors in more than one member of the family.

Gliomas account for most of the malignant neoplasms of the central nervous system and twin studies of these have shown no appreciable concordance. The few case-control comparisons have also shown no significant familial aggregation (Hauge and Harvald, 1958, and Van der Wiel, 1960).

As matters now stand, there is no evidence of geographic or familial aggregation of gliomas and meningiomas. If the lack of such aggregation is not due to the heterogeneity of the primary brain tumor group or the existing relatively crude reporting and collecting systems, we can ponder the causative mechanism that might explain this situation and reflect on the other epidemiologic and genetic approaches which may be indicated.

SUMMARY AND CONCLUSION

Questions on the etiology of several neurological diseases are raised by their patterns of geographic distribution and population selectivity; these have been considered in the light of available genetic information. The even distribution in the United States of the primary brain tumors and the lack of significant familial aggregation of the principal glioma and meningioma groups are difficult to interpret, but no less so than the remarkable north to south variation of multiple sclerosis. The relative roles of genetic and environmental factors for these disorders are still to be determined.

The tremendous research opportunities provided by "geographic isolates" of new or unusually prevalent diseases were reviewed and the fluctuating viewpoint on the role of genetics in amyotrophic lateral sclerosis (ALS) was considered at length. The even geographic distribution and relatively low incidence of ALS in the United States and several other countries contrast with those observed in the Mariana Islands of the western Pacific. Detailed evaluation of the cases has raised some questions about the singularity of sporadic and familial ALS in the United States and of the highly endemic form in the Marianas.

It is hoped that the opportunity presented by the study of the phenomenal incidence of the neurological and other prevalent disorders among the cooperative population of these islands may help to answer the questions which will not only benefit the people of the

islands but also the thousands of persons throughout the world who annually are destined to develop this tragic affliction. If the intensive programs now under way in chronic disease research, as exemplified by the ALS studies, expedite the etiologic solution to any of these problems, the disciplines of geographic pathology and population genetics would be furthered considerably.

REFERENCES

Acheson, E. D. 1961. Multiple Sclerosis in British Commonwealth Countries in the Southern Hemisphere. *Brit. J. Prev. & Soc. Med.* 15: 118–125.

Acheson, E. D., Bachrach, C. and Wright F. 1960. Some Comments on the Relationship of the Distribution of Multiple Sclerosis to Latitute, Solar Radiation and Other Variables. *Acta Psychiat. et Neurolo. Scand.* 35: 132–147.

Alter, M., Allison, R. S., Talbert, O. R. and Kurland, L. T. 1960. Geographic Distribution of Multiple Sclerosis. A Comparison of Prevalence in Charleston County, South Carolina, U.S.A., and Halifax County, Nova Scotia, Canada. *World Neur.* 1: 55–70.

Alter, M., Halpern, L., Kurland, L. T., Bornstein, B., Leibowitz, U., Silberstein, J 1962. Multiple Sclerosis in Israel—Prevalence Among Immigrants and Native Inhabitants. *Arch. Neur.* 7: 253–263.

Aran, F. A. 1850. Recherches sur une maladie non encore decrite du systeme musculaire (atrophie musculaire progressive). *Arch. Gen. de Med.* 84: 172.

Arnold, A., Edgren, D. C. and Palladino, V. S. 1953. Amyotrophic Lateral Sclerosis: Fifty Cases Observed on Guam. *J. Nerv. & Ment. Dis.* 117: 135.

Bondue'le, Bouygues P., Delahousse, J. and Faveret, C. 1959. Evolution simultanée chez un même malade d'une maladie de Parkinson et d'une sclérose latérale amyotrophique. *Revue Neurologique.* 101: 63–66.

Boudin, G. and Barbizet, J. 1956. Formes familiales de sclérose latérale amyotrophique (étude de deux famil'es). *Revue Neurologique.* 95: 229–245.

Brown, M. R. 1951. "Wetherbee Ail": Inheritance of Progressive Muscular Atrophy as Dominant Trait in Two New England Familles. *New Engl. J. Med.* 245: 645–647.

Bunina, T. L., Konowalow, N. V. 1962. Intracellular Inclusions in Hereditary Amyotryphic Lateral Sclerosis, *Korsakov J. of Neuropath. and Psychiat.* 62: 1293–1299.

Cruickshank, E. K., Montgomery, R. D. and Spillane, J. D. 1961. Obscure Neurologic Disorders in Jamaica. *World Neur.* 2: 199.

Fischer, A. and Fischer, J. L. 1961. Culture and Epidemiology: a Theoretical Investigation of Kuru. *J. Health & Hum. Behavior.* 2: 16–25.

Friedman, A. P. and Freedman, D. 1950. Amyotrophic Lateral Sclerosis. *J. Nerv. & Ment. Dis.* 111: 1.

Fullmer, H. M., Siedler, H. D., Krooth, R. S. and Kurland, L. T. 1960. A Cutaneous Disorder of Connective Tissue in Amyotrophic Lateral Sclerosis. A Histochemical Study. *Neurology.* 10: 717–724.

Gajdusek, D. C. 1963. Motor-Neurone Disease in Natives of New Guinea. *New Engl. J. Med.* 266: 474–476.

Gajdusek, D. C. and Zigas, V. 1959. Kuru—Clinical, Pathological & Epidemiological Study of an Acute Progressive Degenerative Disease of the Central Nervous System Among Natives of the Eastern Highlands of New Guinea. *Amer. J. Med.* XXVI: 442–469.

Gervas, G. D. 1959. Une Caso Familiar de Esclerosis Lateral Amiotrofica. *Revista Clinica Española Año.* XX. Tomo LXXIII: 210–212.

Goldberg, I. and Kurland, L. T. 1962. Mortality in 33 Countries from Diseases of the Ne vous System. *World Neur.* 3: 444–465.

Haberlandt, W. F. and Glanville, E. V. 1962. Progress in Neurological Genetics; in *Expanding Goals of Genetics in Psychiatry.* (F. J. Kallmann, ed.) New York & London: Grune and Stratton. 126–135.

Hanau, R. 1960. Su un caso di sclerosi laterale amiotrofica con turbe psichiche. *Studi Sassaresi.* XXXVIII: 119–130.

Hasaerts-Van Geertruyden. 1955. Sur la sclérose latérale amyotrophique héréditaire (une deuxième souche belge). *J. Génét. Hum.* 4: 152–163.

Hauge, M. and Harvald, B. 1958. Genetics of Intracranial Tumors. *Acta Genet.* 7: 573–591.

Hirano, A., Kurland, L. T., Krooth, R. and Lessell, S. 1961. Parkinsonism-Dementia Complex—An Endemic Disease on the Island of Guam. I. Clinical Features. *Brain.* 84: 642–661.

Hirano, A., Malamud, N. and Kurland, L. T. 1961. Parkinsonism-Dementia Complex—An Endemic Disease on the Island of Guam. II. Pathological Features. *Brain.* 84: 662–679.

Hirano, A. and Zimmerman, H. M. 1962. Alzheimer's Neurofibrillary Changes—A Topographic Study. *Arch. Neur.* 7: 227–242.

Kimura, K., Kaneko, Z. and Nogi, K. 1961. Epidemiological and Geomedical Studies on Amyotrophic Lateral Sclerosis and Allied Diseases in the Kii Peninsula (Japan). Presented at the Symposium on Geographic Neurology, Rome, Italy.

Koerner, D. R. 1952. Amyotrophic Lateral Sclerosis on Guam: A Clinical Study and Review of the Literature. *Ann. Int. Med.* 37: 1204.

Konowalow, N. W., Popowa, L. M. and Chondkavian, O. A. 1963. Zur Histopathologie der poliomyelitis anterior subacuta. *Psych. Neurol. und Medizenesche Psych.* 15: 215–226.

Krooth, R., Macklin, M. and Helbish, T. 1961. Diaphyseal Aclasis (Multiple Exostosis) on Guam. *Amer. J. Hum. Genet.* 13: 340.

Kurland, L. T. 1958. Descriptive Epidemiology of Selected Neurologic and Myopathic Disorders with Particular Reference to a Survey in Rochester, Minnesota. *J. Chron. Dis.* 8: 378–418.

Kurland, L. T. 1957. Epidemiologic Investigations of Amyotrophic Lateral Sclerosis. III. A Genetic Interpretation of Incidence and Geographic Distribution. *Proceedings of the Staff Meetings of the Mayo Clinic.* 32: 449–462.

Kurland, L. T., Alter, M. and Bailey, P. 1957. Geomedical and Other Epidemiologic Considerations of Multiple Sclerosis. *Extrait du VI Congrès International de Neurologie.* 11–24.

Kurland, L. T., Faro, S. N. and Siedler, H. 1960. Minamata Disease. *World Neur.* 1: 370–395.

Kurland, L. T., Hirano, A., Malamud, N. and Lessell, S. 1961. Parkinsonism-Dementia Complex—An Endemic Disease on the Island of Guam. Clinical, Pathological, Genetic and Epidemiological Features. *Trans. Amer. Neurol. Ass.* 115–120.

Kurland, L. T. and Mulder, D. W. 1954. Epidemiologic Investigations of Amyotrophic Lateral Sclerosis. I. Preliminary Report on Geographic Distribution, with Special Reference to the Mariana Islands, Including Clinical and Pathological Observations. *Neurology.* 4: 355–378, and 4: 438–443.

Kurland, L. T. and Mulder, D. W. 1955. Epidemiologic Investigations of Amyotrophic Lateral Sclerosis. II. Familial Aggregations Indicative of Dominant Inheritance. *Neurology.* 5: 182–196, and 5: 249–268.

Kurland, L. T., Mulder, D. W., Sayre, G. P., Lambert, E., Hutson, W., Iriarte, L. L. G. and Imus, H. A. 1956. Amyotrophic Lateral Sclerosis in the Mariana Islands (A Scientific Exhibit). *A.M.A. Arch. Neur. & Psych.* 75: 435–441.

Kurland, L. T., Mulder, D. W. and Westlund, W. P. 1955. Multiple Sclerosis and Amyotrophic Lateral Sclerosis: Etiologic Significance of Recent Epidemiologic and Genetic Studies. *N. Engl. J. Med.* 252: 649–653, 697–702.

Kurland, L. T., Myrianthopoulos, N. C. and Lessell, S. 1962. Epidemiologic and Genetic Considerations of Intracranial Neoplasms. Chapter I in *The Biology & Treatment of Intracranial Tumors* by Fields and Sharkey, *et al.* Springfield, Ill.: Charles C. Thomas.

Kurland, L. T. and Reed, D. 1964. Geographic and Climatic Aspects of Multiple Sclerosis. *Amer. J. Public Health.* 54: 588–597.

Kurland, L. T. and Westlund, K. B. 1954. Epidemiologic Factors in the Etiology and Prognosis of Multiple Sclerosis. *Ann. N.Y. Acad. Sci.* 58: 682–701.

Légrand, R., Linguette, M., Delahousse, J. and Gerard, A. 1959. A propos d'un nouveau cas d'association d'une maladie de Parkinson et d'une sclérose laterale amyotrophique. *Revue Neur.* 101: 191–193.

Limburg, C. C. 1950. Multiple Sclerosis and the Demyelinating Diseases; Chapter II in *The Geographic Distribution of Multiple Sclerosis and Its Estimated Prevalence in the United States.* Baltimore, Md.: Williams & Wilkins.

Mackay, R. P. and Myrianthopoulos, N. C. 1958. Multiple Sclerosis in Twins and Their Relatives—Preliminary Report on a Genetic and Clinical Study. *A.M.A. Arch. Neur. & Psychiat.* 80: 667–674.

Mathai, K. V. Feb. 1, 1963. Personal communication.

Mulder, D. W. 1957. *Proceedings of the Staff Meetings of the Mayo Clinic.* 32: 427.

Mulder, D. W. and Kurland, L. T. 1954. Amyotrophic Lateral Sclerosis in Micronesia. *Proceedings of the Staff Meetings of the Mayo Clinic.* 29: 666–670.

Müller, R. 1953. Genetic Aspects of Multiple Sclerosis. *Arch. Neur. & Psychiat.* 70: 733–740.

Myrianthopoulos, N. C. and MacKay, R. P. 1960. Multiple Sclerosis in Twins and Their Relatives: Genetic Analysis of Family Histories. *Acta genet.* 10: 33–47.

O'Brien, W. M., Burch, T. A. and Kurland, L. T., 1964. Hyperuricemia and Gout in the Mariana Islands (unpublished data).

Patrikios, J. 1951. Sclérose latérale amyotrophique avec mouvements involontaires des doigts et du poignet gauches de caractère extra-pyramidal. *Revue Neurologique.* 85: 60–62.

Pratt, R. T. C., Compston, N. D. and McAlpine, D. 1951. Familial Incidence of Disseminated Sclerosis and Its Significance. *Brain.* 74: 191–232.

Reed, D. W., Elizan, T. and Plato, C. 1964. Epidemiologic Features of Amyotrophic Lateral Sclerosis and Parkinsonism-Dementia Complex on Guam. Personal communication.

Robertson, E. E. 1953. Progressive Bulbar Paralysis Showing Heredofamilial Incidence and Intellectual Impairment. *A.M.A. Arch. Neur. & Psychiat.* 69: 197.

Sayre, G. P. 1957. Amyotrophic Lateral Sclerosis; Pathologic Aspects. *Proceedings of the Staff Meetings of the Mayo Clinic.* 32: 447–448.

Sercl, M. and Kovarik, J. 1963. On the Familial Incidence of Amyotrophic Lateral Sclerosis. *Acta Neur. Scand.* 39: 169–176.

Shiraki, H. Nov. 1962. Personal communication.

Stazio, A., Kurland, L. T., Bell, L. G., Saunders, M. and Rogot, E. 1964. Multiple Sclerosis in Winnipeg, Manitoba. Methodological Considerations of Epidemiologic Survey. Ten-year Followup of a Community-wide Survey, and Population Resurvey. *J. Chron. Dis.* 17: 415–438.

Torres, J., Iriarte, L. L. G., and Kurland, L. T. 1957. Amyotrophic Lateral Sclerosis among Guamanians in California. *Calif. Med.* 86: 385–388.

Van der Wiel, H. J. 1960. Inheritance of Glioma. Amsterdam: Elsevier Publishing Co., D. Van Nostrand Co., Ltd.

Van Bogaert, L. and Radermecker, M. A. 1954. XXII. Scléroses latérales amyotrophiques typiques et paralysies agitantes héréditaires, dans une même famille, avec une forme de passage possible entre les deux affections. *Mschr. Psychiat. Neur.* 127: 185–203.

Van Nuis, C. 1964. Age Distribution of Deaths from Amyotrophic Lateral Sclerosis in the United States, 1959–1961. Personal communication.

Williams, G. R., Fischer, A., Fischer, J. L. and Kurland, L. T. 1964. An Evaluation of the Kuru Genetic Hypothesis. *J. Génét. Hum.* 13: 11–21.

Zilber, L. A., Bajdakova, Z. L., Gardasjan, A. N., Konowalow, N. V., Bunina, T. L. and Barabadze, E. M. 1963. Study of the Etiology of Amyotrophic Lateral Sclerosis. *Bull. World Health Org.* 29: 449–456.

PART III

Problems in the Experimental Design of Epidemiological Studies

Selection Techniques for Rare Traits*

By Leslie Kish, Ph. D., *Survey Research Center, The University of Michigan, Ann Arbor, Mich.*

AN OUTLINE OF PROBABILITY SELECTION METHODS

Probability sampling requires that every population element receive a known nonzero probability of selection. In practice this means that the selection of elements be made with a mechanical operation that assures the desired probabilities, generally the proper use of a good table of random numbers. The role of probability selection in general research strategy has been discussed elsewhere (e.g. Kish, 1959). The virtues of probability sampling, and methods for attaining it have been discussed in books (Cochran, 1963; Deming, 1960; Hansen, Hurwitz and Madow, 1953; Yates, 1960; Kish, 1965), and in many articles (e.g. Kish, 1953). The review in this section is necessarily brief to the point of distortion.

There exist different procedures for selecting a sample of n from a finite population of N elements. Unrestricted random sampling with replacement is seldom practiced on finite populations. For sampling without replacement the basic selection method is *simple random sampling.* It admits with equal probability each of the C_n^N possible combinations. Other selection procedures may be viewed as modifications introduced to provide more precise, economical and practical designs. They introduce controls of various kinds that eliminate some of the undesirable combinations among the C_n^N, thus increasing the probabilities of the more desirable combinations. Generally, the more desirable combinations either reduce the element cost, or decrease the element variance.

The major kinds of modifications are shown in table 1, with the alternatives available for each. The alternatives can occur in com-

*Writing supported by Grant USPHS/GM09889-02 from the National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare.

bination with each other, and this permits a large variety of possible designs. The variety increases rapidly in multistage samples, where the design must be decided for each successive stage of selecting sampling units. The introduction of different estimates, utilizing auxiliary variables, further increases the variety of possible sample designs. Finally, variations in the practical aspects of field and office procedures add to the flexibility of selection designs, and permits their adaption to diverse situations.

TABLE 1.—*A Taxonomy of Probability Selection Procedures*

a. Epsem	Unequal probabilities.
b. Unstratified	Stratified sampling, optimum allocation.
c. Element sampling	Cluster sampling; subsampling; multistage selection.
d. Random choice	Systematic selection.
e. One-phase selection	Two (multi-) phase sampling.

a. *Epsem* stands for "equal probability of selection method," and describes any sample in which all population elements receive equal selection probabilities. Simple random sampling is only one kind of epsem, and epsem is a special case of probability sampling. Epsem is used widely, because it leads to "self-weighting" means, where the sample mean is a good estimate of the population mean. Epsem can also be obtained from unequal selection probabilities that are made to compensate each other through the several stages of multistage selection. In contrast to epsem, some designs call for unequal probabilities for different groups of elements; this "loading" or over-sampling is generally compensated with inverse weights in the statistics.

b. In *stratified sampling* the population is divided into strata, so that every element appears in one stratum. The sample is selected independently within each stratum. The stratum estimates are then weighted and combined into the population estimate. Thus the population mean is estimated with

$$\bar{y}_w = \sum_h W_h \bar{y}_h, \ (h=1, 2, \ldots, H), \qquad (1)$$

where $\bar{y}_h$ is the mean of the sample within the h th stratum, and usually $\bar{y}_h = \sum_i y_{ht}/n_h$. The combined mean $\bar{y}_w$ has the variance

$$\text{var}\ (\bar{y}_w) = \sum_h W_h^2 \ \text{var}\ (\bar{y}_h). \qquad (2)$$

This is the weighted sum of the variances, var$(\bar{y}_h)$, within each stratum. The known weights represent the relative sizes of the strata, and $\sum_h W_h = 1$. Most often the strata are measured in numbers of population elements, and $W_h = N_h/N$, where N_h denotes the numbers of

elements in the h th stratum. If within each stratum a simple random sample of elements was selected then (2) becomes

$$\text{var}\,(\bar{y}_w)=\frac{1}{N^2}\sum_h N_h^2\,(1-f_h)\,\frac{s_h^2}{n_h}, \tag{3}$$

where $s_h^2=\sum_i (y_{hi}-\bar{y}_h)^2/(n_h-1)$ is the element variance computed from the stratum sample. The $f_h=n_h/N_h$ denotes the sampling fraction within the stratum, and the "finite population correction," $(1-f_h)$, is often negligible.

The most widely known selection method is proportionate sampling of elements, when the sampling fraction is the same in all strata: $f=n/N=n_h/N_h$. Then the sample proportions represent exactly the population proportions: $n_h/n=N_h/N=W_h$; and the sample is "self-weighting."

On the contrary, disproportionate selection rates may be introduced deliberately, to aim at the "optimum allocation" of the sample sizes, n_h, assigned to the strata. Suppose that a significant portion of the cost can be expressed as $C=\sum_h n_h J_h$, where J_h is the cost per element in the h th stratum. Then C is minimized for a specified variance, or the variance is minimized for a specified C, when the sampling fractions f_h in the strata are made directly proportional to S_h, the element standard deviations, and inversely proportional to $\sqrt{J_h}$:

$$f_h=\frac{n_h}{N_h}=K\,\frac{S_h}{\sqrt{J_h}} \tag{4}$$

where K is a constant of proportionality. The unequal selection rates are compensated with inverse weights for the sample cases.

Disproportionate allocation leads to large gains when we can separate strata in which the S_h or the J_h are very different, and known at least roughly. For example, suppose that we need to estimate the proportion $\overline{M}$ of a class with a rare trait in the population. Perhaps we can isolate small strata that contain large portions of the population; this means finding strata in which the proportion's $\overline{M}_h$ with the rare trait is much greater than in other strata. Then the element standard deviations are $S_h=\sqrt{\overline{M}_h(1-\overline{M}_h)}$ for the binomial variable representing the proportions $\overline{M}_h$ of the class in the several strata. Since $S_h \doteq \sqrt{\overline{M}_h}$, we can say that the optimum rates will be approximately $f_h=K\sqrt{\overline{M}_h/J_h}$. For optimum allocation to be worth the trouble, one must be able to locate differences among the $\overline{M}_h/J_h$, with factors of 6, 10 or 100. This can seldom occur on a geographic basis; but special lists and registrations often may qualify.

Instead of estimating the proportion $\overline{M}$ of the rare class, we may need the mean of a trait Y_1 based on the rare class who possess the trait, with $Y_1 \neq 0$; and disregarding the large portion $(1-\overline{M})$ of the population without the trait, for whom $Y_1=0$. The problem arises when one cannot separate beforehand the rare class, whose members appear as a random subclass in the sample of each stratum. It received explicit treatment (Yates, 1960, p. 300; Kish, 1961; Kish, 1965, section 4.5), on which we base the following conclusions. The mean for a subclass from a stratified sample can be computed as

$$\bar{y}_w = \frac{\sum_h y_h/f_h}{\sum_h m_h/f_h} = \sum_h w_h \bar{y}_h, \qquad (5)$$

where $y_h = \sum_i^{m_h} y_{hi}$ is the total of the Y_1 variable obtained on the m_h class members, found in the sample of n_h cases from the h-th stratum. Further, $\bar{y}_h = y_h/m_h$, the sample mean in the stratum; $\overline{m}_h = m_h/n_h$, the proportion of the class; and

$$w_h = \frac{m_h/f_h}{\sum_h m_h/f_h},$$

the stratum weights of the class found in the sample. If the population numbers M_h of the class were known constants, then we would have the known weights W_h instead of the variates w_h; and instead of the special forms (5) and (6) we could use (1) and (2).

The variance of $\bar{y}_w$ can be found to a good approximation to be

$$\text{var}\,(\bar{y}_w) \doteq \sum_h (1-f_h)\frac{w_h^2}{m_h'}[t_h^2 - \overline{m}_h(\bar{y}_h - \bar{y}_w)^2], \qquad (6)$$

where $m_h' = m_h(n_h-1)/n_h$ and $t_h^2 = \sum_h (y_{hi}-\bar{y}_w)^2/m_h$.

Examination of (6) discloses that for small subclasses, when the $\overline{m}_h(\bar{y}_h-\bar{y}_w)^2$ are small, the gains of proportionate stratification vanish. But the gains of optimum allocation remain. This allocation can be shown to be reached approximately when $f_h = k\sqrt{\overline{M}_h t_h^2/J_h}$. When the t_h are approximately equal then the optimum rates are $f_h = k'\sqrt{\overline{M}_h/J_h}$, where k' is a constant of proportionality.

Finally the population aggregate Y of the Y_1 variable in a subclass can be estimated with ordinary formulas, by also including all the $Y_1=0$ values of nonmembers of the class (Hansen, 1953, p. 149–155). But it can also be estimated from the class members alone as

$$\tilde{Y} = \sum_h y_h/f_h, \qquad (7)$$

with the variance

$$\text{var}\,(\tilde{\bar{Y}}) = \sum_h (1-f_h)\,\frac{N_h^2}{n_h-1}\,\overline{m}_h(t_h^2+\bar{y}_w^2-\overline{m}_h\bar{y}_h^2)\cdot \qquad (8)$$

c. *Cluster sampling* denotes selection methods in which the sampling units are clusters of elements. Contrariwise, in element sampling the elements are also the unique sampling units. Examples of cluster sampling are selecting entire families or entire city blocks to sample the persons within them; or selecting entire classes or entire schools to observe their students; or selecting entire villages to survey their inhabitants.

Just what is a convenient cluster depends on the survey situation and resources, and these are matters of practical expediency. First, the elements are defined in accord with the research aims. Then the sampler must decide whether he can retain the simplicity of element sampling, or must he abandon it for some convenient clusters that permit a more economical and practical design.

Cluster sampling can be useful when it decreases considerably the element cost, by lowering the costs of travel and listing. But it often increases the element variance, and it also complicates the analysis. It should be used when the lowered element cost more than compensates for the increase in element variance and in the cost of the analysis. The larger the average cluster the greater are these effects. When the effects are too great we may search for an economic compromise between the two conflicting effects of clustering. This search leads to subsampling or two-stage sampling; similar considerations can lead to more stages and to multi-stage sampling.

Two further complications are present typically in clustered designs. First, clusters are generally selected after stratification, because this is both feasible and desirable. Second, in natural situations one must generally deal with clusters of unequal size. Even when the selection design creates clusters of roughly equal sizes, their exact equality is destroyed by nonresponse, by imperfect size measures, and by the fluctuations in the sizes of subclasses on which the statistics are often based.

Despite the possible variations in complex designs, basic formulas exist, that can be used rather generally for the sample mean and its variance. They also hold for means of subclasses. The sample mean is a *ratio mean*, in which both numerator and denominator are random variates. If from each stratum two primary selections (*a* and *b*) are chosen at random in the first stage of selection, the ratio mean may be denoted as

$$\frac{\bar{y}}{\bar{x}} = \frac{\sum_h y_h}{\sum_h x_h} = \frac{\sum_h (y_{ha}+y_{hb})}{\sum_h (x_{ha}+x_{hb})}\cdot \qquad (9)$$

735-712 O—65——12

Here y_{ha} and y_{hb} are the sample sums of the Y_i variable for the two primary clusters selected independently at random from the h th stratum. Any needed weighting is incorporated in them so that the expected value for both y_{ha} and y_{hb} is (except for a constant) Y_h, the population aggregate of the Y_i variable in the h th stratum. For epsem selections the y_{ha} and y_{hb} may be the simple sample sums. Similar statements hold for the x_{ha} and the x_{hb}, which usually represent merely the element counts in the primary selections. The variance of the ratio mean is approximately

$$\text{var}\left(\frac{y}{x}\right) \doteq \frac{1-f}{x^2}\left[\sum_h D_{y_h}^2 + \left(\frac{y}{x}\right)^2 \sum_h D_{x_h}^2 - 2\left(\frac{y}{x}\right)\sum_h D_{y_h} D_{x_h}\right], \quad (10)$$

where

$$D_{y_h}^2 = (y_{ha}-y_{hb})^2,\ D_{x_h}^2 = (x_{ha}-x_{hb})^2, \text{ and } D_{y_h} D_{x_h} = (y_{ha}-y_{hb})(x_{ha}-x_{hb}).$$

When the number of primary selections in the stratum is a_h, we can have more generally

$$\frac{y}{x} = \sum_h \sum_{a_h} y_{ha} \Big/ \sum_h \sum_{a_h} x_{ha} \quad (9')$$

with the variance (10) now defined by $D_{y_h}^2 = \left(a_h \sum_{a_h} y_{ha}^2 - y_h^2\right)/(a_h - 1)$, and

similarly for the other two terms. These formulas can be found justified and derived in several places (Kish, 1959; Kish, 1965, section 6.5).

Clustering often has a drastic effect on the variance. A sample of a clusters appears confined, compared to an element sample which is spread over n points of the population. The distribution of variables within natural clusters is typically not random, but more or less homogeneous. The elements of clusters tend to be similar; for example, people in blocks, students in classes tend to resemble each other. This effect of clustering on the element variance can be represented approximately as $[1+rho(n/a-1)]$. Here *rho* is the coefficient of intraclass correlation; it is generally positive in natural clusters; it is small for most variables, and its maximum value is $+1$. Hence the clustering effect is not great if the mean cluster size, n/a, is small. When n represents the sample total of a rare class, then n/a can be kept small, even if we use clusters rather large in terms of their complete count (Hansen *et al.*, 1953, pp. 107–109).

d. *Systematic selection*, an alternative to random choice, denotes the selection of sampling units with an interval of k applied to listings after a random start from 1 to k. In subsampling it can be a practical substitute for random choices.

e. *Two-phase sampling* refers to subselecting the final sample from a larger sample that was preselected to provide information for

improving the final selection. The first phase can be a screening operation that cheaply separates small strata containing high proportions of the rare target population.

SELECTION TECHNIQUES FOR RARE TRAITS

A small proportion $\overline{M}=M/N$ of a population of N elements is characterized by a trait denoted with variable $X_i \neq 0$, whereas for the vast majority $(N-M)$ the variable $X_i=0$. The search for elements that are rare but important can have any of three aims, but the problems and methods of selection are similar for all three. (a) We may need to know only M or $\overline{M}$, the number or proportion of the members of the class with the trait; for example, the prevalence of a rare genetic trait. Estimating M and $\overline{M}$ represent the same problem if the population base, N, is known well enough relative to M. (b) We may need $\overline{Y}=Y/M=\overline{X}/\overline{M}$, the mean of the X_i variable within the class, disregarding the majority $(N-M)$ for which $X_i=0$. Here $Y=X$ and $\overline{Y}$ denote the sum and the mean of the X_i in the class. For example, X_i may denote the individual's medical costs caused by the trait and $\overline{Y}$ the mean cost in the class. (c) Or we may need $\overline{X}=X/N=\overline{M}\overline{Y}+(1-\overline{M})0$, the mean in the entire population; for example, the mean cost of a trait in the population. Of course, $\overline{M}$ is only the special case of $\overline{X}$ when $Y_i=1$, the mere presence of the variable for all M class members. Estimating $\overline{X}$ from $\overline{Y}$ is easy when $\overline{M}$ is known well enough relative to $\overline{Y}$.

The degree of rarity and the cost of measuring the trait are the two factors that mainly determine the relative worth of different selection methods, as noted in a previous discussion (Cochran, 1954). Sample surveys are used widely and successfully to obtain cheap measurements of traits that are not too rare. A screening procedure of the population may be practical to find a class if $\overline{M}$ comprises 10 or 20 percent of it, but much less so if $\overline{M}$ is 1 in 100 or 10,000. Even difficult measurements, if not too rare, have already been applied with probability sampling, as for fertility behavior (Freedman *et al.*, 1959), and medical examinations (U.S. National Health Survey, 1964). On the contrary, very rare traits are easy to sample only if finding them is cheap. An extreme example would be a centralized and complete list of persons with a well recognized rare disease.

But rare traits that are expensive to find present difficult sampling problems. Yet they may be undertaken with some of the following techniques. These may be used in combinations, and relations between them may be noted. In choosing between them consider the trait's rarity. Consider also the available survey resources, including knowledge about large concentrations of the trait. This could mean

lists, or known geographical concentrations, or high correlations with other clear indicators.

a. *Multipurpose samples.* Several rare traits can be discovered in one sample, and its costs can be divided among the cooperating investigations. This approach can also yield data about relationships between the traits. Some economic and market surveys are based on such cooperation. The National Health Survey may be considered a combination of many specialized purposes.

b. *Cumulation* of a rare population can be obtained as byproduct of continuing surveys with changing samples that cover large populations in time. The results must be interpreted as an average over the period covered by the surveys, but this may offer little difficulty. First, past surveys may already contain the data in an acceptable uniform manner; though too rare in single samples, the accumulation may provide an adequate base for analysis. Second, the measurement can be added to future surveys if the enumerators can perform it. Even a lengthy addition adds little to the average interview length, if applied only to a small portion; but emphasis is needed to keep enumerators alert to rare events. Third, the surveys may perform only a screening operation to identify the rare individuals; the difficult measurements on them is done separately. If the identification is imperfect it can still provide the first of a two-phase sample.

c. *Large clusters* can decrease drastically the costs of locating and screening population elements. For example, a sample of complete blocks of a city, or of entire villages of a country, can be searched (screened) for persons with the trait. For samples of the entire population large clusters are avoided, because of their adverse effects on the variance; but they will yield only small clusters of the rare trait if this is widely spread. If, on the contrary, it is concentrated in small areas, these can be often recognized and stratified.

d. *Controlled selection.* Sometimes the search has been confined to one or two hospitals (or other places), purposively chosen. A broader base can be secured with four or eight replication, chosen to represent a wide spectrum of backgrounds (Kish, 1965, section 14.4). But perhaps the sample can be expanded to 15, 30, or 60 hospitals (places). Then one prefers to control carefully the selection to make it represent the population along several important control variables. The number of control cells may far outrun the number of permitted selections. The selection problem has been treated under "controls beyond stratification," "two-way stratification," "lattice sampling," and "controlled selection" (Hansen, *et al.*, 1953; Yates, 1960; Hess, 1961; Kish, 1965, section 12.8).

e. *Disproportionate stratified sampling.* It is efficient to increase the sampling fractions from small strata if these contain large portions of the rare trait, demarcated from large portions of the total population containing only small portions of it. In other words, if

most of the rare trait can be located in small pockets of the total population it should be sampled heavily. If, for example, 90 percent of the rare trait can be located within 10 percent of the total population, then large gains can be made. But geographical strata can seldom yield high enough separations of class densities. We look for special lists of the class with the needed trait, or with another trait highly correlated with it.

However, only small gains can be had from the presence of only one of the two desirable conditions: either from high concentrations of only a small portion, or from mild concentrations of a large part of the population. This can be seen in the formulas for "optimum allocation" of the sample sizes into strata, yet it is often misunderstood. For example, U.S. cities have exclusive neighborhoods that are full of rich people; but oversampling them will not bring large gains, because most rich people live scattered elsewhere. Similarly, visitors to a set of clinics may contain a high concentration of the class with the trait, yet most of the class may never reach any clinic. Sometimes the visitors to the clinics may well represent the entire class; but often the two kinds of class members may differ drastically. For example, rich people in exclusive neighborhoods differ in many ways from those living outside them. An exotic example is found in interstellar space: Although almost empty compared to stellar density, it contains nevertheless a large portion of total matter, and its composition may differ from stellar matter.

A high degree of separation is needed, because optimum allocation for a rare class typically calls for selection rates about proportional to $\sqrt{M_h/J_h}$ where M_h is the proportion of the rare class, and J_h is the element cost in the stratum. Large differences in M_h are needed to obtain even moderate differences in $\sqrt{M_h}$.

f. *Two-phase sampling* (double sampling) may be needed to distinguish, in the first phase of preliminary screening, two or more strata in which the proportions of the rare elements differ greatly. Then the methods of disproportionate sampling can be applied in the second phase. The first phase is needed when desirable stratifying variables are not readily available for the entire population. The allocation of effort, between the first phase (screening) sample and the second phase, depends on the utility of the screening and on the relative costs of the two phases.

g. *Screening* operations underlie several methods. They are useful when we find a cheap screening test that has very few *false negatives* (erroneous exclusions), and not many *false positives* (erroneous inclusions). Applied to the population, or to a large sample in two-phase sampling, it picks out a smaller sample of screened inclusions, the potential members of the rare class. This part of the sample is then measured with the more precise and accepted standard process to discard the false positives.

What about false negatives? Even good screening can be imperfect. In the absence of overwhelming assurance of safety, a sample check of the screened exclusions is advisable on large surveys. If this finds none of the rares, it provides the error's upper limit. But if false exclusions rise beyond a negligible proportion, then the screened exclusions become a stratum to be sampled at a lower rate.

The nature of screening operations vary: a brief interview, or rapid observation by field worker, or scanning data on filed cards. Generally the screening should err more on the side of false positives than of false negatives. The unit cost for the former (some extra work) is less than for the latter (either bias or reduced efficiency). The costs and the proportions of both kinds of error should enter the design of the screening process.

h. *Special lists* provide, when available, the best means for locating a rare target class. If they are very rare, widely scattered, and difficult to identify, a list may be the only means for finding them. The ideal list contains the entire target class (or a good sample) clearly isolated or identified, and with the information (e.g., address) needed to make the survey.

Such lists can sometimes be produced with ingenuity, perhaps with tolerable redefinitions of the class, and perhaps by combining several lists. The economics of the joint use of several selection frames has been investigated recently by Hartley (1962).

Snowball sampling is the colorful name (Goodman, 1961) for techniques of building up a list of a special population, or a sample, by using an initial set of its members as informants. An example given by Deming (1963) is asking an initial selection of deaf people to supply the names of other deaf persons they knew. It may be difficult to assign reliable measures of the probability of inclusion. However, members of some populations may know about each other, as the deaf do in towns and cities that are not too large; perhaps also owners of pigeons or of racing cars.

Extra effort is needed if the list contains impurities, foreign elements and replicate listings. But the worst problems are caused by incompleteness, the *noncoverage* of elements missing from the list (Kish, 1965, sections 2.7, 11.1–11.5, and 13.3). When the list is good yet imperfect we must decide whether to go after the small missing portion or to omit it. If the missing portion is small and extremely difficult to find we may be forced to accept the "cutoff" bias. But if needed and feasible, it may be included with a supplemental procedure in a special stratum. The decision depends on several considerations.

First, the element cost is high when the missing units are rare, widely scattered and difficult to find. Hence, their inclusion should be designed with a considerably lower sampling fraction. Second, the bias of their exclusion should be judged against the other sources

of errors, hence also against the size of the sample. The importance of a fixed bias increases with the precision and size of the sample. The size of the bias should be balanced against the standard error, and the cost of reducing the bias against the reduction in the standard error it can buy.

Third, the effect of exclusion is different for estimating frequency totals and for mean values of a trait. The exclusion of missing values has little effect on the mean, but great effect on the total, if the missing values resemble closely the values found on the special list. Contrariwise, if the missing values are of negligible magnitude they will affect the mean more than the aggregate. Hence one should consider both the number and the nature of the missing units. In that judgment he may be guided by knowledge based on theory, and on past investigations. If necessary he may conduct a pilot study to assess at least roughly the limits of the problem.

i. *Supplements* are useful for sampling the portion missing from the special list, when that portion is small but not negligible. If the missing portion is negligible we need only the special list; if it is large we don't need the list at all. Consider the rare elements as coming from two strata: the special list to which the sampling fraction f_a is applied, and the supplement for the rest of the population sampled with the fraction f_b. Because the element cost is typically much higher in the supplement, the sampling rate f_b should be considerably lower, according to the optimum allocation of resources.

The separation of the two strata (A and B) is complete if the supplement B excludes any element that appears on the entire special list A, whether or not it appears in the sample from that list. In that case the elements in the two samples should receive weights proporional to $1/f_a$ and $1/f_b$.

But the exclusion of A elements from the B sample may be inconvenient. Then they may also be permitted to receive the selection probability f_b in the supplement, additional to their selection probability f_a on the special list. It is necessary and sufficient that all the A elements in the entire sample, from both the A and B samples, be identified. Then all these must receive weights proportional to $1/f_a'$, because $f_a' = (f_a + f_b)$ was their probability of selection if the two samples were independent. If duplicate appearances in the sample are excluded, than $f_a' = f_a + f_b - f_a f_b$.

SUMMARY

In summary, a *special list* that includes the rare class and nothing else is best, especially for very rare traits. If the list is good but imperfect, a *supplemental sample* may find the missing units. Lacking a good list we try to separate a small *special stratum* that contains

most of the rare class, from a large stratum that contains much less. The special list or stratum is sampled with *disproportionately high rates*. If we can't find such special strata we try to create them with a cheap *screening* operation. The efficiency of the screening may be increased with *double sampling*, and by using *large clusters*. The costs of screening may be shared in *multipurpose* surveys, or the rare class may be *accumulated* in a periodic survey operation.

REFERENCES

1. Cochran, W. G., 1963: Sampling Techniques, New York. *John Wiley and Sons, 2d ed.*
2. Cochran, W. G., 1954: Problems Confronting Human Genetics: Discussion. *American Journal of Human Genetics* 6: 77–80.
3. Deming, W. E., 1960: Sample Design in Business Research, New York. *John Wiley and Sons.*
4. Deming, W. E., 1963: Chapter 1 in Rainer, J. B., Altshuler, K. Z., and Kallman, F. J. (Eds.): Family and Mental Health Problems in a Deaf Population, New York. *New York State Psychiatric Institute, Columbia University.*
5. Freedman, R., Whelpton, P. K., and Campbell, A. A., 1959: Family Planning, Sterility and Population Growth, New York. *McGraw-Hill.*
6. Goodman, L. A., 1961: Snowball Sampling. *Annals of Mathematical Statistics* 32: 148–170.
7. Hartley, H. O., 1962: Multiple Frame Surveys. *Proceedings of the Social Statistics Section, American Statistical Association,* 203–206.
8. Hansen, M. H., Hurwitz, W. N., and Madow, W. G., 1953: Sample Survey Methods and Theory, 2 vols., New York. *John Wiley and Sons.*
9. Hess, I., Riedel, D. C., and Fitzpatrick, T. B., 1961: Probability Sampling of Hospitals and Patients, Ann Arbor, Michigan. *Bureau of Hospital Administration.*
10. Kish, L., 1953. Selection of the Sample Book. Chapter 5 in Festinger and Katz (eds.): Research Methods in the Behavioral Sciences, New York. *Dryden Press,* 175–240.
11. Kish, L. and Hess, I., 1959: On Variances of Ratios and Their Differences in Multi-Stage Samples. *Journal of the American Statistical Association* 54: 416–446.
12. Kish, L., 1959: Some Statistical Problems in Research Design. *American Sociological Review* 24: 328–338. Also in the *Bobbs-Merrill Reprint Series in the Social Sciences.*
13. Kish, L., 1961: Efficient Allocation of a Multi-Purpose Sample. *Econometrica* 29: 363–385.
14. Kish, L., 1965: Survey Sampling, New York. *John Wiley and Sons.*
15. U.S. National Health Survey, 1964: Cycle I of the Health Examination Survey. *National Center for Health Statistics,* Washington, *Series* 11, No. 1.
16. Yates, F., 1960: Sampling Methods for Censuses and Surveys, New York. *Hafner, 3rd ed.*

The Reliability of Survey Data

By Charles F. Cannell, Ph. D., *Survey Research Center, University of Michigan, Ann Arbor, Mich.*

The question of the reliability of surveys can be approached by means of three propositions:

1. Some data collected in some surveys have a high level of accuracy.
2. The same data collected in other surveys are highly inaccurate.
3. Some data collected in any surveys are always inaccurate.

To be symmetrical there should be a fourth proposition: Some data collected in any surveys are accurate. But this proposition is not included since even the simplest data show inaccuracies.

As the propositions indicate, the major components of survey accuracy are the type of information sought and the methods used to collect it. (I am not including in this paper any considerations of sampling since these were covered by Dr. Kish.)

The purpose of this paper is to examine some of the factors relating to characteristics of the information and of data collection methods, and to consider some of the variables which contribute significantly to the accuracy or inaccuracy of surveys.

We turn first to a consideration of some of the characteristics of the data themselves and the degree to which the wanted information is accessible to the person who is asked to report it. Clearly, the first requisite to accurate reporting is that the respondent have these data in his possession.

One point is so well known that we need spend little time on it. A person is likely to have more information about himself than he has about others. For example, in a study of the validity of the reporting of hospital episodes conducted for the National Health Survey (Health Statistics, Series D, No. 4), we found that when an individual was asked to report about his own hospitalization the under-reporting rate was 7 percent. When reporting for close relatives the under-reporting was somewhat worse, and rose to 22 percent when the

person being reported for was a more distant relative. Several factors are involved in this under-reporting rate, one of the most obvious being that in many cases the individual asked to report never knew the information and hence could not report it correctly.

But of more importance is the inaccessiblity of information when the researcher and the respondent do not share the same concepts or even the same language. This problem is illustrated by a questionnaire administered to farmers, which contained the following question: "In planning your farming operations, do you use inductive or deductive thinking?" This is an extreme example of the lack of shared concepts, in which the objective may be perfectly clear to the researcher but incomprehensible to the respondent. Problems of this type are so apparent that we avoid them almost without thinking. But less extreme examples are not so easily avoided. Frequently we would like to ask respondents about such diseases as diabetes, hypertension, or arthritis. Take the question, "Have you had rheumatic fever?" Asked of a sample of respondents, such a question will receive either "yes" or "no" responses, with only a few reporting that they do not know. A "yes" response may have any of several meanings. It may mean, "I have been under treatment for a condition which was diagnosed by a doctor as rheumatic fever," or, "I have something like my brother had and he said he had rheumatic fever so I think I have it," or, "I guess I must have it, but I just call it plain rheumatism." Conversely, a "no" response may mean, "I never had anything I knew to be rheumatic fever," or, "I know what rheumatic fever is and the doctor told me I didn't have it," or, "I don't know what it is so I guess I don't have it."

The lack of shared concepts between the researcher and the respondent is one of the major limitations of the survey technique as a method of providing data sought by the medical researcher. Parenthetically we might note that some surveys have been severely criticized for inaccuracy because the information obtained from patients in interviews as to conditions from which they suffered did not check with diagnostic data obtained by direct medical examination. To indict surveys on this basis is fallacious. The fault lies not with the surveys but with the researchers who use the survey method improperly.

The potentialities of survey research as a method of collecting data are limited to information which the individual possesses and to concepts which he understands. Unfortunately, information which the respondent thinks he has is often incomplete or inaccurate. A study conducted by the Health Insurance Plan of Greater New York under contract with the National Health Center (Health Statistics, Series D, No. 5) compared diagnostic material reported by respondents with material in doctors' records. Major discrepancies in many categories between the records and the reports of the respondents were revealed. Our study of hospitalization data (Health Statistics, Series D, No. 4),

which compared hospital records with reports of patients, showed that respondents reported the diagnoses of malignant neoplasms at only 75 percent of the rate of the hospital records, whereas benign or unspecified neoplasms were reported at 150 percent. We suspect that much of this error represents not only the failure of the doctor to transmit the correct diagnosis to the patient, but also represents a perceptual distortion on the part of the respondent as to his own condition.

The doctor is the major source of the respondent's information about his physical condition. Obviously respondents cannot report information which doctors have not given them, or information which has been given them inaccurately, or information which has been given accurately but has been misunderstood. One of the primary difficulties in getting respondents to report what physicians have told them is caused by the esoteric language which physicians use. For example, one person reported to us that his doctor told him he had "hemorrhoids in the rectory."

Although the doctor is the main source of information which the respondent has about his physical condition, unfortunately for researchers he is not the only source. Friends, relatives, and one's own information frequently lead to self-diagnoses. These diagnoses, reported to the interviewer as readily as the more authoritative doctor's diagnosis, cannot be expected to be as accurate. Recently we discovered a new version of this problem in a health survey while interviewing an elderly Polish couple. The wife reported that she had not visited a doctor in 20 years because she was much afraid of doctors. Yet, when asked about her physical condition, she gave a couple of specific diagnoses unlike those commonly made by laymen. The puzzled interviewer asked her how she had arrived at these diagnoses, and she explained that her husband, who had a chronic condition, went to the doctor every other week. When she had a particular set of symptoms she related them to her husband, who reported them to the doctor as his own symptoms. The doctor then made a diagnosis and gave him medicine which he took home for his wife.

Thus far we have concentrated on what one *cannot* expect from surveys. Returning to a positive approach, we may ask what a respondent can report and report accurately. First of all, the respondent can tell us how he feels. He can tell us any degree of disability he may have, whether he is able to work, whether he has been confined to bed because of illness, whether his work has been restricted, and so forth. He can report specific symptoms and can describe these symptoms in detail. He can report the frequency or duration of symptoms, their intensity, and what he does to relieve them. He can report what he considers to be the cause of the symptoms, that is, what he considers to be the diagnosis. He can go further and relate

the facts surrounding the diagnosis, whether it was made by himself, by his friends, or by a medical authority. He can give us the history of his conditions, illnesses, and injuries. The respondent can report these types of information plus many other types he has about himself and under certain conditions will report accurately. It is relevant to point out that many of these data cannot be obtained from any records whatever but must come from the respondent himself.

To recapitulate: Surveys are potentially accurate or inaccurate depending upon the uses to which they are put. A major obstacle to accurate information is that the respondent may not have the data which the researcher wishes to obtain. Surveys cannot be expected to replicate data from medical records any more than medical records can be expected to provide data on undiagnosed conditions for people who have never sought medical aid. For example, if one wanted to know how many people have been hospitalized for particular diseases, surveys would be of little use. It would be better to obtain the information from hospital records. But on the other hand, if one wanted to know how many people had respiratory infections during a particular week, records would not help. A survey approach asking about symptoms and degree of disability would be the only source of information since only a small proportion of those persons would have been to a doctor. If one wanted to know how many people have hypertension, the only accurate procedure would probably be to examine a cross-section sample of the population, a technique which the Health Examination Survey of the National Health Survey is using currently.

Assuming that the survey approach is used and that the respondent has access to the information wanted, the important question then arises: Will the information be reported accurately? There are two main obstacles to accurate reporting: memory, the ability to recall accurately, and motivation, the willingness of the respondent to report accurately. The problem of memory is a familiar one. In the study of hospitalization referred to earlier, it was found that the rate of reporting was high for episodes which occurred close to the date of the interview, but dropped consistently and with increasing rapidity the longer the elapsed time between the interview and the date of the hospitalization. For episodes occurring a year prior to the interview, the rate of under-reporting of episodes was several times the rate for the first few weeks immediately preceding the interview. If the event in which the researcher is interested is both recent and from the respondent's point of view significant, a simple question may be enough to bring the episode fully into his memory. But if much time has elapsed since the event, the problem is more difficult. The gradual decay of information is manifested in many ways. Events of trivial significance for the respondent may be forgotten almost as quickly as they occur. Even experiences which were once prominent

are likely to be forgotten if they have little relevance to the individual's current life. The adult is unlikely to remember the age at which he developed chicken pox, even though the event may have had dramatic importance for him as a child. For such routine matters as one's breakfast menu or the content of the previous evening's television programs, recollection may be gone within a matter of hours.

Memory is a complex function of elapsed time since the event, current cues or relevance for present affairs, and the original significance of the event to the individual. The usual effect of these processes is a reduction in the amount and accuracy of information available to the researcher. In the extreme case, the reduction is complete, the respondent is unable to recall the event. More often, however, the reduction is partial, the respondent can recall that he had a particular experience but is unable to describe it in accurate detail. Thus Goddard *et al.* (1961) have found that although mothers' evaluations of difficulty of labor and delivery agreed with the physicians' records, information on the length of gestation period, feeding problems of the infant, and illnesses during infancy showed marked distortions.

Unfortunately for the researcher, the process of memory decay is not uniform and orderly. The events of the past do not fade gradually from view while retaining their original dimensions. On the contrary, the process of forgetting and remembering involves considerable distortion. Certain aspects drop from view, others are elevated to prominence, and the entire past is recalled in a way that makes it plausible and consistent in terms of the individual's present experience. Related to this selective process of memory is the distortion of previous experience in the general directions of social acceptability or preservation of one's self-image. Wenar (1963), for example, reports that when a mother distorts developmental facts about her child she tends to do so in a way that makes her child appear more precocious; and that when she distorts child-training practices, she tends to bring them in line with Dr. Spock.

Frequently, such distortions are not conscious falsification of information but result from unconscious repression of facts which are intolerable to the self. Studies of political behavior provide an interesting example of this process. In national samples voters have been asked repeatedly to name the candidate for whom they voted in the previous presidential elections. The longer the election recedes into the past, the greater the proportion of people who report that they voted for the winning candidate.

When the individual is asked to report some past event, a number of complex processes occur simultaneously, and in combination they affect and even determine the availability of the information. These processes have to do with the initial importance of the event and its meaning for the respondent; they have to do with the tendency toward

congruence and plausibility of previous experience; and they relate to various ego defense mechanisms.

We now come to respondent motivation, that is, the willingness of the respondent to report those things which are accessible to him and are not subject to the distortion of memory. Respondent motivation is coming to be recognized as perhaps the most important factor in determining the amount and the accuracy of the data available to the researcher. Thus, hospital episodes involving arthritis, deliveries, hernias, appendicitis, and gall bladder disease have shown less than 5 percent under-reporting, while episodes involving mental or personality disorders have shown under-reporting of 32 percent (Health Statistics, Series D, No. 5). In general, episodes classified as non-threatening, that is, as not embarrassing or detrimental to the individual's self-image, showed an under-reporting rate of 10 percent in this study of hospitalization data. Episodes rated as threatening were under-reported 21 percent. Further, for the episodes classified as threatening which were reported, the report of the diagnosis was altered to make it more palatable or more acceptable.

Other fields are replete with examples of this type. In economic studies, individuals with small savings accounts tend to overstate their size, while people with particularly large accounts tend to understate them. Men in high income groups are less likely to admit having borrowed from a cash lender than those in the lower income groups.

All of these data indicate that obtaining accurate responses requires more than merely approaching a respondent with a set of questions. The researcher, recognizing the great value of his research, somehow expects the respondent to be equally eager to make the study a success. Yet, for many studies, there is little reason apparent to the respondent for divulging information to the interviewer. The interviewer's request for information does not mean that the respondent shares the researcher's goals even to the extent of being willing to think carefully about each question.

The interview is a complex interaction of forces. The respondent has an image of himself as a particular kind of person; he has a set of social norms which tell him what behavior is appropriate and what is inappropriate. Thus, the mother may be reluctant to admit to particular kinds of illnesses on the part of her child because, to her, it implies that she has in some way been a poor mother. The husband may not be willing to report that he has been sick or has had to stay home in bed, because his image of himself is that of a healthy, self-reliant individual. Any admission of illness may be, to him, a sign of weakness or dependency. A person may be quite willing to report his appendectomy but, because of social norms, most reluctant to talk about his venereal disease.

In some cases, however, the problem of motivation is simpler than this. The individual can report recent conditions which he has suffered but may still be unwilling to expend the effort to ensure the accuracy of dates and other relevant information. The reason for such inaccuracy is that the respondent may not share the researcher's goals or appreciate the relevance of his participation in the interview. This problem can be solved by generating forces strong enough to overcome the negative factors. Fortunately, much of the information in which the epidemiologist is interested does not involve forces which are so negative that the task of establishing strong positive motives is impossible.

You may feel that I have presented a dismal picture of the potentialities of surveys as a useful device in epidemiological studies. You may think further that this is an odd point of view for a person whose main activity during the past several years has been in survey research. I took this approach because survey research techniques are deceptively simple. Too many people have criticized surveys because they have used them improperly. All too often they have decided to do a study, sat down and thought up a few questions, sent out a couple of laboratory assistants to take interviews, and then have been surprised that the data they collect are not accurate. Then they tend to condemn unjustly survey research in general. My purpose has been to point out the importance of understanding the variables of reporting, and to stress the importance of using adequate techniques. Survey research is a method of measurement. As such it has the strengths and weaknesses of most measuring instruments. It also has appropriate and inappropriate applications. One can expect to get some things but not others from it. Since survey research is a relatively new technique, there is much about the methodology which we do not know as yet.

In conclusion, I would like to emphasize the need for studies of methods and techniques designed to make survey research more valuable to epidemiology. The National Health Survey of the Public Health Service has an active program of methodological investigations, and is to be commended for this attention. But more is needed.

There is need also for a facility for collating and evaluating the experience of researchers who have used survey methods so that others may have the benefit of their successes and failures. Minimally this means that each research report based on survey data should contain the questions and instruments used, as well as information on the types of interviewers employed and the techniques they used. Eventually, and hopefully not too far in the future, one can envisage a facility for data retrieval to which researchers can turn for information on the methods and questions found most successful in obtaining data needed for a particular objective.

It is through research on methods and sharing of experience that survey research methods will become more accurate and will become more useful in epidemiological studies.

REFERENCES

1. Goddard, Katharine E., Broder, G., and Wenar, C., 1961. Reliability of pediatric histories, a preliminary study. Pediatrics *28:* 1011–1018.
2. Health Statistics, Series D, No. 4, 1961. Reporting of hospitalization in the health interview survey. U.S. Department of Health, Education, and Welfare.
3. Health Statistics, Series D, No. 5, 1961. Health interview responses compared with medical records. U.S. Department of Health, Education, and Welfare.
4. Wenar, Charles, 1963. The reliability of developmental histories—summary and evaluation of evidence. Unpublished mimeographed report, University of Pennsylvania School of Medicine.

A Trial of Methods for Collecting Household Morbidity Data*

By Kenneth R. Wilcox, Jr., M.D., Dr. P.H.,
Division of Laboratories, Michigan Department of Health, Lansing, Mich.

So far this morning, Dr. Kish and Dr. Cannell have presented some of the methods and problems of sampling and surveys. I would like to present a specific example of some of these methods and problems. This small survey was primarily designed to compare three methods for collecting reports of acute illnesses. The purpose of presenting such a study in a meeting that is primarily concerned with chronic disease is to illustrate some of the aspects of household surveys that should be considered each time this method of collecting information is used. Although some of the aspects presented are primarily related to acute illness surveys, some are common to most types of surveys.

The study reported here was a pilot study to aid in the development of a morbidity surveillance system in Tecumseh, Michigan, as part of the Tecumseh Community Health Study. The staff of the Community Study furnished invaluable assistance in the planning and execution of the study. The surveillance itself is intended to supplement the extensive, detailed information on chronic diseases which has been collected for the past few years and which is still being collected. Since the survey method developed was to be used in a continuous surveillance of subsamples of the community, some basic requirements were imposed on it. The method chosen must be fairly simple and quick to use so that a reasonably large sample of the community might be reached each week. On the other hand, it must aim at a broad spectrum of illnesses so that unusual occur-

* This study was done as partial fulfillment of the requirements for the degree of Doctor of Public Health from the University of Michigan, School of Public Health. The study was supported by grants G-26 and H-6378 from the National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare.

735-712 O—65——13

rences of unexpected types of illness are not missed. The method must also obtain some basic information about the illnesses reported.

Two household interviews and one daily record system were used in the study. The interviews were based on two different frames of reference for illnesses. The questions used in Interview A were based on modifications of activity that might be produced by illness as shown in figure 1. The names of the household members were placed in the spaces across the top of the form and Y or N circled in the columns below the names for the appropriate answers to all questions. A check was placed in the space below the names of the household member or members who served as respondents. Two general questions were asked about illnesses and accidents occurring during the prior four weeks for each member of the family individually. Then a series of questions were asked about "anyone in the family" concerning illnesses that caused them to go to bed, stay in the house, miss work or school, just feel "under the weather," or go to a hospital. In addition, questions were asked about medicines taken or physicians

Form A-2: Interview A

NAME OF HOUSEHOLD MEMBER Mark R for respondent in box below name						
1. Would you tell me about any times that ____(you or person's name) were sick since about_____?	Y	N	Y	N	Y	N
2. Did___(you, he, she) have any accidents or did___hurt___self in any way since_____? For example, did_have any cuts, bruises, sprains or other injuries?	Y	N	Y	N	Y	N
3. Were there any (other) times when anyone in the family had to go to bed because they were sick or did not feel well since_____?	Y	N	Y	N	Y	N
4. Were there any (other) times when anyone had to stay in the house because they were sick or did not feel well since_____?	Y	N	Y	N	Y	N
5. Did anyone have to miss work or school any (other) times (since_____?)	Y	N	Y	N	Y	N
6. Were there any (other) times when anyone did not feel well or was under the weather but still could get around (since_____?)	Y	N	Y	N	Y	N
7. Did anyone take any (other) medicines or drugs (since_____), such as aspirin, antacids, home remedies or drugs prescribed by a doctor? (If YES, see other side)	Y	N	Y	N	Y	N
8. Did anyone talk to a doctor or see a doctor for any (other) reason (since_____?) (If YES, see over)	Y	N	Y	N	Y	N
9. Did anyone (else) have to go to the hospital (since_____?)	Y	N	Y	N	Y	N

FIGURE 1

visited in an attempt to elicit reports of other illnesses not yet mentioned. The data about all illnesses reported were recorded on a separate form.

The second approach to interviewing was based on a list of symptoms and diseases modified from the work of Cartwright (Cartwright, 1959). Figure 2 shows a portion of the two sides of the form con-

Form A-3: Interview B

These cards have some common symptoms and sicknesses written on them. Would you take the cards and look at them one at a time. If you or any other person in the family has been bothered by the condition on the card since about _____, would you tell me.

Mark R for respondent in box below name

1.	Aches or pains in muscles or other parts of the body	Y	N	Y	N
2.	Backaches	Y	N	Y	N
3.	Unusual bleeding from any part of the body	Y	N	Y	N
4.	Breathlessness or shortness of breath	Y	N	Y	N
5.	Convulsions, seizures, or epilepsy attacks	Y	N	Y	N
6.	Diarrhea	Y	N	Y	N
7.	Dizziness or faintness	Y	N	Y	
8.	Earache or ear trouble	Y	N	Y	N
9.	Eye trouble or trouble seeing	Y	N	Y	N
10.	Fever	Y	N	Y	N
11.	Female trouble	Y	N	Y	N
12.	Hay fever or asthma	Y	N	Y	N
13.	Headache	Y	N	Y	N
14.	Heart trouble or heart palpitations	Y	N	Y	N

FIGURE 2

Form A-3 (Continued)

	NAME			
15.	Indigestion, pain in stomach or stomach ache	Y	N	Y
16.	Infectious diseases or childhood diseases	Y	N	Y
17.	Itching, skin rash or skin sores	Y	N	Y
18.	Kidney trouble or trouble passing water	Y	N	Y
19.	Pain in chest	Y	N	Y
20.	Pain or swelling in any joints	Y	N	Y
21.	Rectal trouble	Y	N	Y
22.	Rheumatism or arthritis	Y	N	Y
23.	Runny nose, cold, sore throat or cough	Y	N	Y
24.	Sick to stomach or vomiting	Y	N	Y
25.	Sleeplessness or unusual tiredness	Y	N	Y
26.	Swelling of ankles	Y	N	Y
27.	Trouble with teeth, gums or mouth	Y	N	Y
28.	Unusual nervousness or irritability	Y	N	Y
29.	Weakness of any part of the body	Y	N	Y
30.	Accidents or injuries such as cuts, burns, sprains	Y	N	Y
31.	Has anyone used any medicines or taken any shots	Y	N	Y

ASK: Has anyone had any other illnesses or problems with their health that you haven't mentioned yet? _____

COMMENTS:

taining the list of items. The responses were recorded on this form in the same manner as on the Interview A form. Each item was typed on a 3-inch by 5-inch file card and the deck of cards was presented to the respondent at the time of the interview. As the respondent went through the cards one at a time, she indicated which member of the family had had the symptom or disease during the prior 4 weeks. If the respondent went through the cards too fast, the interviewer could ask those items, for which responses were not recorded, from the printed list. In some cases, more than one symptom might contribute to a single illness. The illnesses elicited were entered on a separate form and the required data recorded.

The daily record is shown in figure 3. The content of the form generally follows Interview A and is divided into four parts. The first three parts were answered with checks to indicate: 1) the occurrence of an illness or accident, 2) any changes in activities associated with such events, and 3) whether or not any medicines had been used or a doctor seen. The fourth part contained space for description of the illness or accident that occurred. Each sheet was printed on legal size paper and contained sufficient space for 1 week's entries for one household member. A sufficient supply of sheets to last 4 weeks for all of the family members was stapled in a folder, and a sheet of instructions stapled to the inside cover. The daily records were placed in the homes by interviewers who explained their use to the household respondent. One week later the interviewer returned to answer any further questions, and at the end of the period the diaries were picked up. The daily record used in this study is similar to the one used in the California Health Survey study in San Jose, California (Allen *et al.*, 1954).

The common form used to collect information about illnesses obtained from both interviews and the daily records is shown in figure 4. The two parts of this form were printed on two facing pages. Each illness was recorded on a separate line across both pages. The nature of the illness, the time of onset and cessation, consultation with a physician, and various types of disability were determined. The illnesses were coded and data processing cards were punched directly from these forms. A different arrangement appeared at the bottom of the form for recording accidents.

The primary purpose of the study was to compare the ability of these three procedures to obtain reports of illnesses. The numbers of reported illnesses, or rates derived from these numbers, were used for the comparison. It was assumed that, for these acute illnesses, the method obtaining the most reports was the best. Although it might have been valuable to investigate the contributions of the various component parts of the procedures and the effect of varying their order of presentation, such a study was beyond the scope of this survey.

Form A-5: Daily record

TECUMSEH COMMUNITY HEALTH STUDY-DAILY HEALTH RECORD

FOR (Name of person): _____ Week of:

	EACH DAY, check (X) any of these three columns which apply:			EACH DAY, check (X) the columns below which apply:						EACH DAY, check (X) in these columns if:		EACH DAY, If any columns except 3 and 9 have been checked, describe what the trouble was:
				Didn't carry on usual activities--had to:				Carried on usual activities--				
	Was sick, had an ailment or did not feel as well as usual	Had an accident or injury or felt effects of previous accident or injury	Felt as well as usual	Stay in hospital	Stay in bed at home	Stay in the house but not in bed	Miss work or school	But not able to do as well as usual	As well as usual	Any medicines or drugs were taken	A doctor was seen or contracted	(a) For each sickness write in what was wrong. If a doctor was seen, tell what he said the trouble was. (b) For each accident or injury, be sure to tell what parts of the body were hurt, how each part was hurt, and how and where the accident happened. (c) If any medicines were taken, tell what the medicine was and what it was taken for. If any illness stays the same for more than one day, write in "Same". (Use the back of this sheet if more room is needed.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mon.												
Tue.												
Wed.												
Thu.												
Fri.												
Sat.												
Sun.												

If in hospital as a patient this week, give name and address of hospital:

Give the name of any doctor seen during this week:

FIGURE 3

TCHS-3/62

In an attempt to facilitate the comparison between the three procedures, replicates of three families each were selected from the Tecumseh population by matching on several family characteristics, such as, family size, number of children in the family, the age and education of the household head, the geographic location of the household within the city, etc. Each family in a matched group or replicate was randomly assigned to one of the three procedures. A total of 50 replicates were chosen so that 50 families would be contacted with each procedure. Families who had moved from the area or could not be reached were replaced from a list of alternates that had been prepared before the field work began.

Four weeks was chosen to be the period over which information would be collected. The arrangement of the procedures and family groups is shown in table 1. At the end of the first 4-week period, both family groups were interviewed and the daily records placed. At the end of the second period, the first two family groups were interviewed again with the alternate procedure, and the daily records were retrieved. All of the interviews could not be done in 1 week so that there was some overlap in the actual weeks included in each period for different families. In the discussion of the results of the study, only the incidence of illnesses occurring during the 4-week periods, that is, those illnesses having their onsets within the period will be used.

TABLE 1.—*Arrangement of family groups and data-collecting procedures during two 4-week periods*

		Family group		
		I	II	III
Period	1	Interview A	Interview B	----------
	2	Interview B	Interview A	Daily record (C)

Before discussing the comparisons and analyses of the data, I would like to present the overall rates and types of illnesses experienced during the study. The rates for reported illnesses for the three groups of families is shown in table 2. The total incidence was approximately one-half of an illness per person per 4-week period. Note that the incidence of reported illness on the daily records was approximately twice that for either of the interviews.

The illness rates obtained in the National Health Survey were compared with ours in order to determine whether or not the procedures used were as adequate as a standard method. The rate for those illnesses reported in our study that caused disability or a visit to a physician was expanded to obtain a rate for 13 weeks as shown in table 3. This rate was similar to the rates reported by the National

_____ page	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24

Name Sex Age (1)	What was wrong?, or Would you mind telling me about it? or How did it bother ____? (2)	When did this ___ (illness, attack, condition) first bother ___? (3)	Does it still bother ___? IF NO When did it stop bothering ___? How often did it bother you since _____? (4)	Did ___ talk to or see a doctor about it? IF YES, What doctor was it? (5)

25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
M F ____																		Y N				Y N			

What did the doctor say the trouble was?... did he give it a medical name?	Did it cause any change in _____ daily activities?				
	IF NO, STOP IF YES: In what way?	Did ___ have to miss work (school)? How long?	Did ___ have to stay in the house? How long?	Did ___ have to stay in bed IF NO, STOP IF YES: How long?	Did ___ have to stay in the hospital? How long? Which hospital?
(6)	(7)	(8)	(9)	(10)	(11)
51 52 53 54 55 56 57 58	59	60 61 62	63 64 65	66 67 68	69 70 71 72 73
	Y N	Y N	Y N	Y N	Y N
	Y				

FIGURE 4.—The Collecting Form

Health Survey for the same time of year for the 4 years prior to this study.

TABLE 2.—*Incidence rates for all families in the study for nonrecurrent illnesses having their onsets within the study period*

Family group	Number of illnesses	Number of person-periods	Rate*
I	227	**453	50.1
II	213	***449	47.7
III	208	223	93.7
Total	648	1125	57.8

*Illnesses per 100 person-periods of observation.
**219 persons during the first period and 234 during the second.
***217 persons during the first period and 232 during the second.

TABLE 3.—*Comparison of the present study to the U.S. National Health Survey by rates of illness causing disability or a visit to a doctor during a 3-month period*

Present study	March 15 through May		
	1962	62.1*	
		January through March	April through June
U.S. National Health Survey.	1958	74.3	48.6
	1959	68.9	54.1
	1960	74.5	41.3
	1961	63.0	45.2

*Rate adjusted from 4-week to 13-week period. Rates are illness per 100 person-periods.

The types of illness reported were approximately the same as would be expected at this time of year, as shown in table 4. About one-third of the illnesses were respiratory, and about one-fifth were accidents. Gastrointestinal complaints were about one-tenth of the total illnesses. The specific infectious diseases contributed only a small proportion of the total illnesses.

TABLE 4.—*Rates for acute illnesses by category of illness*

Category of illness	Number of illnesses	Percent of total number	Rate*
Respiratory	230	35.5	20.5
Accidents	127	19.6	11.3
Gastrointestinal	74	11.4	6.6
Infectious	37	5.7	3.3
Other	180	27.8	15.9
Total	648	100.0	57.8

*Illnesses per 100 person-periods.

The effect of the reported illnesses on activities is shown in table 5. Only one-third of the illnesses caused any change in the person's activities, and only one-fifth of the illnesses were attended by a

physician. Of those illnesses causing some change in activity, about 40 percent caused the person to miss work or school or stay in bed, and about two-thirds of them caused the person to stay in the house when he or she would ordinarily be going out. It is clear that most of the illnesses reported were very minor.

TABLE 5.—*The distribution of acute illnesses by the amount of disability caused*

	Yes		No	No data or not significant*	Total
	Number	Percent	Number		
Illness caused person to:					
Change activities	220	34.0	416	12	648
Miss work or school	87	39.5	85	48	220
Stay in the house	144	65.5	71	5	220
Stay in bed	84	38.2	132	4	220
See a physician	127	19.6	516	5	648

*Not significant applies to those situations where the loss could not apply, e.g., a housewife could not miss work or school.

The primary statistical comparison of the three procedures was done using all of the illnesses reported during the second period from the 48 replicates of families in which all three procedures were done. As shown in table 6, the daily record produced by far the most reports of illnesses. There was a significant difference between the three procedures when tested by both the analysis of variance using replicate families and Friedman's nonparametric analysis of variance (Siegel, 1956). Both methods of analysis were used because the data showed significant departure from normality and inequality of variance, even after logarithmic transformation. A significantly higher illness rate was obtained with the daily record than with either interview by the multiple range test of Duncan. Thus, in terms of total illnesses reported, the daily record was superior.

TABLE 6.—*Comparison of the rates for acute illnesses reported during the second 4-week period for the three procedures used in the study*

Family group	Form used	Number of illnesses	Number of persons in group	Number of illnesses per 100 persons
I	B	116	225	51.6
II	A	83	227	36.6
III	C	208	223	93.3
Total		407	675	60.3

In order to compare the two interviews more adequately, the 42 replicate pairs of families who were interviewed during both times with both procedures were analyzed, as shown in table 7. The rates for the two family groups were about the same. The rate for the first period was higher than that for the second period, which would be expected in the spring when the incidence of illness is generally declining. The rate for Interview B is higher than for Interview A. The

analysis of variance using a replicated Latin square design showed the difference between the two interview procedures to be of only borderline significance ($0.1>P>0.05$), (Snedecor, 1956), and a Wilcoxson matched-pairs signed rank nonparametric test on the replicate families was not significant (Siegel, 1956). This finding would indicate that if the use of this list of symptoms and diseases is the better method to elicit more illnesses, this study was not of sufficient size to demonstrate it.

TABLE 7.—*Comparison of the rates for acute illnesses reported during both 4-week periods for the two interview procedures*

Family group	Number of persons	Time period				Total rate
		1		2		
		Form	Rate*	Form	Rate	
I	205	A	48.8	B	47.8	48.3
II	201	B	60.2	A	41.3	50.7
Total	406		54.4		44.6	49.5

*All rates expressed as the number of illnesses per hundred persons.
Total for form A: 45.1.
Total for form B: 53.9.

The remainder of the analysis was performed to determine the qualitative aspects of the differences between the procedures used. Some analysis was also done to examine certain problems associated with surveys for illness.

The changes in activity associated with the illnesses reported by the three procedures is shown in table 8. In this table, the number of illnesses reported during the second period on the daily records can be compared directly to the number of illnesses reported on the interviews. As noted above, the total number is considerably higher for the daily record. The number causing some change in activity is also higher, but this is probably due to a difference in the meaning of "change in activity" in the daily record as compared to the interviews. For the various categories of changes in activities, the three procedures obtained about the same number of illnesses, except that the daily

TABLE 8.—*Changes in activity associated with the acute illnesses reported with the three procedures for the second period*

Family group	II	I	III
Form used	A	B	C
Number of persons	227	225	223
Total number of acute illnesses	83	116	208
Number of illnesses causing:			
Change in activities	24	29	79
Missed work or school	14	16	17
Confined to house	19	23	25
Confined to bed	13	16	8
Visit to doctor	28	23	28

record obtained fewer reports of persons confined to bed. Clearly, the daily record's main value is for the collection of reports of minor illnesses.

The difference in reporting of illnesses causing changes in activity with the two interviews was determined by comparing the number of illnesses reported during both periods for both interviews, as shown in table 9. Although a greater total number of illnesses were reported with Interview B, more illnesses causing various types of changes in activity were reported with Interview A. This can be explained by the fact that the orientation of Interview A concerned changes in activity associated with illnesses.

TABLE 9.—*Changes in activity associated with the acute illnesses reported with the two interview procedures for both periods*

Family group Form used Number of persons	Both A 406	Both B 406
Total number of acute illnesses	183	219
Number of illnesses causing:		
Changes in activities	74	52
Missed work or school	34	29
Confined to house	64	43
Confined to bed	37	31
Visit to doctor	54	38

Breakdown of the number of illnesses by category of illness is shown in table 10. As might be expected, the daily record produced more reports of illness in all categories except the specific infectious diseases. For both time periods, more illnesses were reported with Interview B for the four categories representing what would generally be considered "illness." This is probably due to the fact that Interview B asked specifically about various illnesses and symptoms without regard to severity. The differences in the reporting of specific infectious diseases probably represents real differences in incidence because these illnesses, most of which were measles, present a striking departure from normal health. By contrast to the general illnesses, more accidents were reported with Interview A during the second period. Accidents were asked about specifically for each individual at the

TABLE 10.—*The number of illnesses reported by each of the three procedures for different categories of acute illness*

Period	1		2			Total
Form	A	B	A	B	C	
Family group	I	II	II	I	III	
Number of persons	219	217	232	234	223	1125
Total number of illnesses	109	125	88	118	208	648
Respiratory	48	53	26	46	57	230
Gastrointestinal	10	18	4	15	27	74
Musculoskeletal, mental, neurological, and special sense organs	8	20	12	24	48	112
Other	6	11	14	14	23	68
Infectious	18	5	8	6	0	37
Accidents	19	18	24	13	53	127

beginning of Interview A; they were asked about "any member of the family" at the end of Interview B. Again, the position and amount of emphasis placed upon the question seemed to affect the amount of reporting for these disorders.

This study is too small to allow generalization, but the results seemed to emphasize the point that the manner in which the questions are asked and the emphasis of the questions may markedly influence the reporting of events. If the study utilizing interviewing is primarily focused on the impact of illness on people's activities, questions about changes in activities should be used. If, on the other hand, the emphasis of the study is on the complete spectrum of illness, the questions used should probe into the symptoms and illnesses by name. A good example of the effect of changing the emphasis is Dr. Cannell's study on hospitalization in conjunction with the National Health Survey (Cannell, 1963). Marked improvement in the under-reporting of hospitalizations was obtained when the intensity of the questioning about these events was increased.

In this study a household respondent was used for convenience. Since respondents usually do not report as completely for others as they do for themselves, the rates of illness reported on the interviews for self-respondents and respondents by proxy were compared as shown in table 11. With Interview A the rate of illness reported is about the same for respondents by proxy as for self-respondents. For Interview B the illness rate is higher for self-respondents than respondents by proxy. The reason for this difference cannot be ascertained from this small study, but two factors may be responsible: 1) asking the two general questions specifically about each member of the household might result in better reporting for others with Interview A; or 2) reporting of the respondent's own minor illnesses is probably better than his reporting of other family members' minor illnesses. It was observed in the Washtenaw County Accident Study that respondents reported more minor injuries for themselves than for others, but for major injuries the reverse was true (Velz and Hemphill, 1953). It was also shown that the rates were higher for those who responded jointly than for those who responded for themselves alone. The whole problem of obtaining information about other members of a household from a single respondent needs more study (Feldman, 1960).

TABLE 11.—*Rates of acute illness reported for self-respondents and respondents by proxy for both procedures and both periods*

Form	Respondent	Period 1	Period 2	Both periods
A	Self	45.1	43.8	44.6
	Proxy	51.2	40.8	47.2
B	Self	74.1	56.1	66.4
	Proxy	52.5	53.1	52.7

One of the problems associated with interviewing for illnesses is the failure of the respondent to report events because he does not remember them. One would expect illnesses occurring yesterday to be reported with more accuracy than those occurring 3 to 4 weeks ago. It is important, however, to obtain some estimate of the magnitude of this loss in reporting with the passage of time. One method which is useful over short periods of time is the comparison of the number of illnesses reported for the week prior to interview with the number reported in weeks more remote from the interview. Such a comparison for these data using all of the interviews is shown in figure 5. The number of illnesses reported declines quite rapidly and the number reported for the fourth week prior to interview is about one-third of those reported for the week prior to interview. A loss of reporting of about the same magnitude has been reported in two other studies as shown in table 12. These studies done at different times and places and using different interview procedures show a remarkably similar loss in reporting for the four weeks prior to interview (Downes and Mertz, 1953; Nisselson and Woolsey, 1959).

TABLE 12.—*The decrease in reporting of illnesses with longer periods of recall in this study and two others*

Weeks prior to interview	Ratio of illnesses reported in prior weeks to illnesses reported for the week just prior to interview		
	Present study*	Westchester County, N.Y.	California Health Survey
Second	0.80	0.75	0.73
Third	.45	.57	.51
Fourth	.39	.37	.42

*Includes both interview procedures and both periods.

This loss in reporting is not the same for all types or severity of illness. For example, hospitalizations are reported almost completely for a 4-week period (Cannell, 1961). The difference between the total illnesses reported and those illnesses causing some change in activity or a visit to a physician are shown in figure 6. During the first period, a rise in the number of illnesses during the second week prior to interview was caused by a brief epidemic of respiratory and gastrointestinal illness and the occurrence of several cases of measles in the population under survey. The number of illnesses for which a physician was visited was not influenced by this increase. During the second period the number of illnesses causing a change in activities or a visit to a physician did not vary from week to week, although the total number of illnesses shows the expected decline with increasing interval to interview.

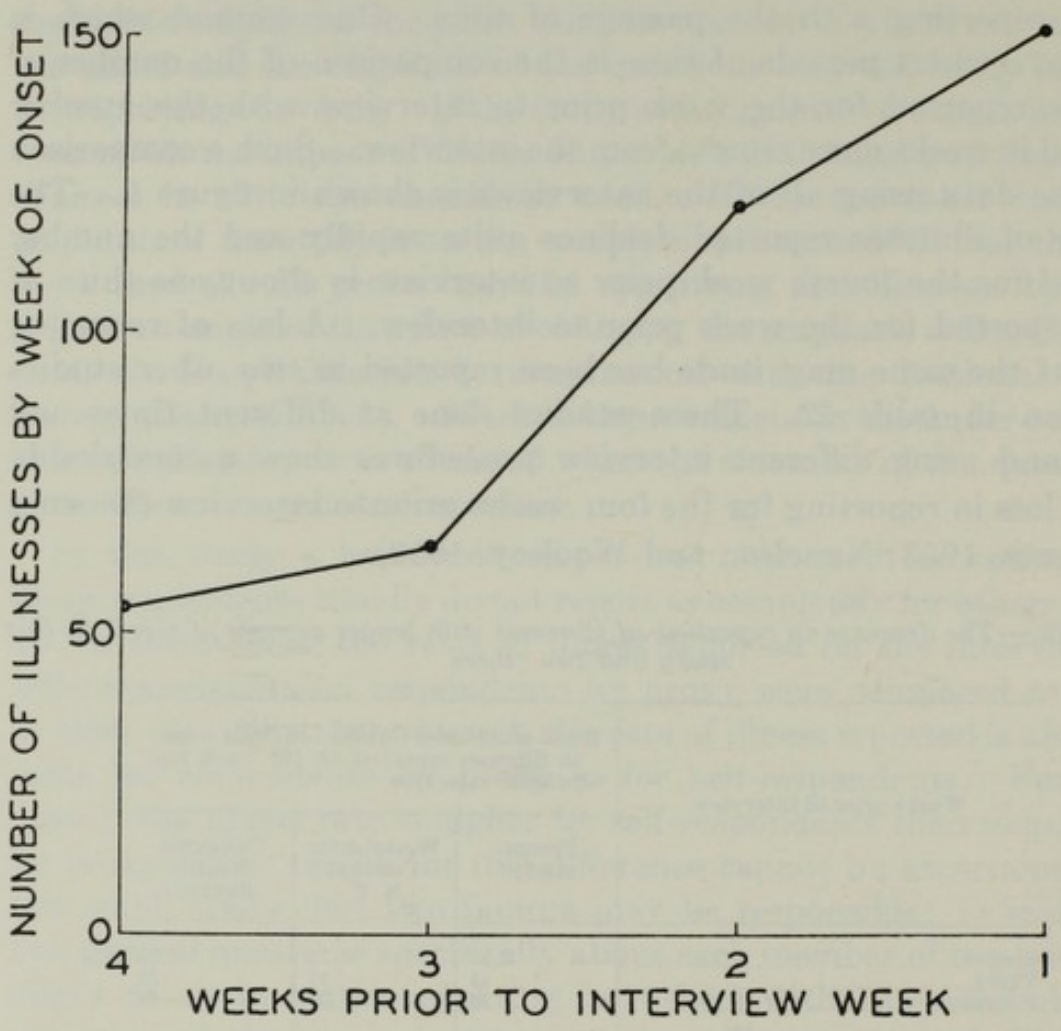


FIGURE 5

The relationship between the rates of illnesses for different weeks reported in the daily record was quite different than that reported in the interviews during the same time period, as shown in figure 7. The high rate observed during the first week of recording was probably due to initial enthusiasm and was followed by a drop in rate to the same level as that obtained by interviews for the week prior to the visit. The daily record gives more complete and consistent reporting of minor illnesses, but other problems connected with the maintenance of such records would have to be solved. Some people forget to keep the records and others do not enter enough information to allow a clear determination of the nature of the illness. Daily records will probably work best in studies of the complete spectrum of a limited group of diseases. For example, such records might be a good method to study acute respiratory illnesses.

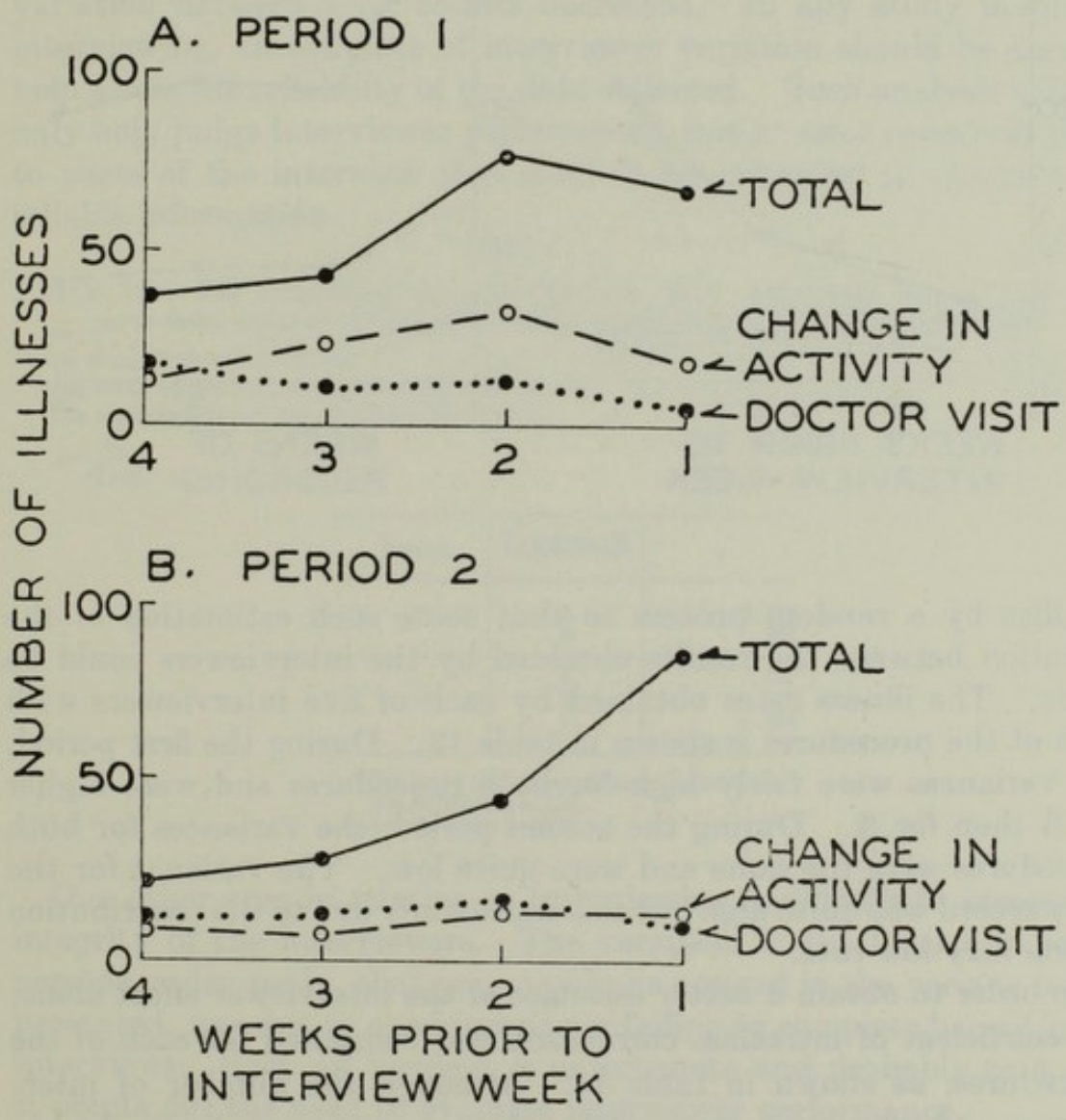


FIGURE 6

One very important determinant of the success of a procedure is the performance of the persons actually doing the work, in this case, the interviewers. A method of assessing this effect on the results of procedures used is to examine the amount of variability between interviewers. In this study, the interviewers were assigned to the

735-712 O—65——14

NUMBER OF ILLNESSES BY WEEK OF ONSET IN RELATION TO THE WEEK OF INTERVIEW. COMPARISON OF THE INTERVIEWS AND THE DAILY RECORD.

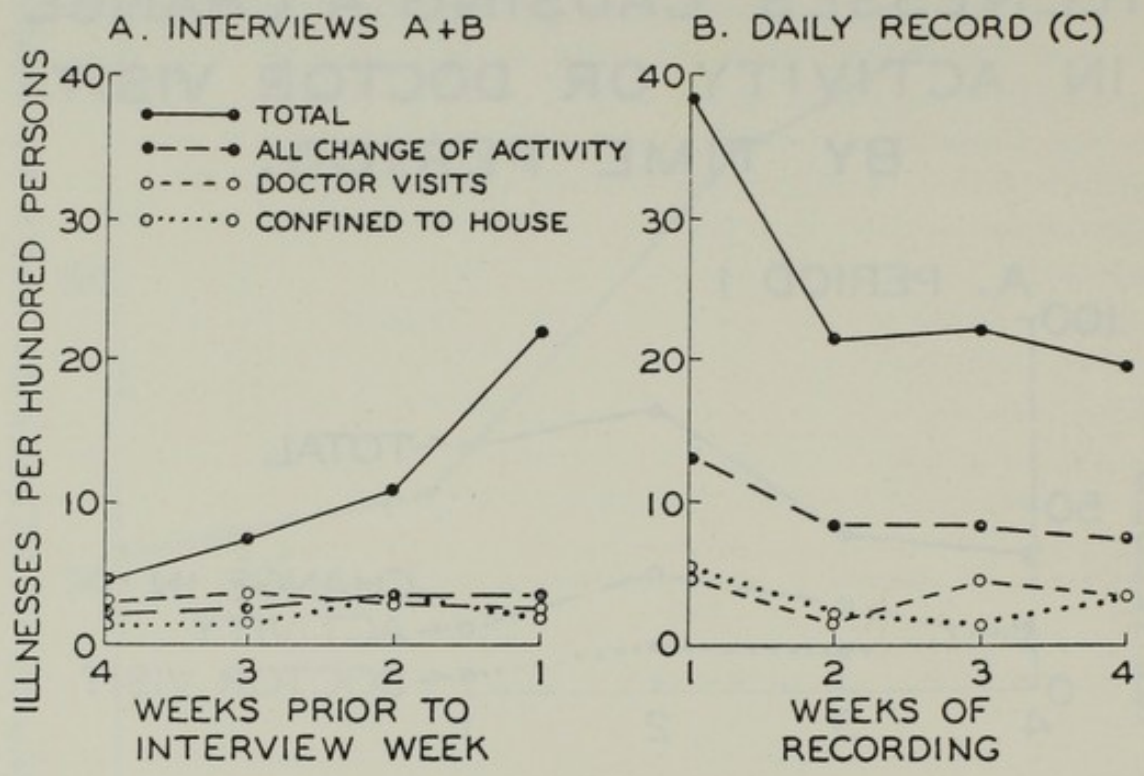


FIGURE 7

families by a random process so that some such estimation of the variation between the results obtained by the interviewers could be made. The illness rates obtained by each of five interviewers with each of the procedures is shown in table 13. During the first period, the variances were fairly high for both procedures and were higher for B than for A. During the second period, the variances for both procedures were the same and were quite low. The variance for the daily record was quite high but this was mainly due to the contribution of one very low rate.

In order to obtain a better estimate of the interviewer effect alone, the coefficient of intraclass correlation was computed for each of the procedures, as shown in table 14. Although the number of interviewers and interviews was small and the significance levels for *rho* were quite high, the *rho* values for the interviews done during the first period were clearly above the significance level and those for the second period were very low. The value of *rho* for the daily records was also low, which might be expected with this type of self-administered record. This would seem to indicate that, as the interviewers became more experienced with the interviewing procedure used, the

TABLE 13.—*Rates of illness reported by the families interviewed by different interviewers*

Interviewer	Procedure-Time period					Overall rate
	A-1	B-1	A-2	B-2	C-2	
a	0.348	0.139	0.342	0.375	1.056	0.461
b	.349	.710	.263	.419	.368	.420
c	.476	.400	.314	.524	1.026	.552
d	.500	.833	.643	.553	1.125	.735
e	.865	.881	.333	.757	1.158	.791
Average	.508	.593	.379	.526	.947	.592
Variance	.0448	.0994	.0227	.0221	.1074	.028

variation between their results decreased. In any study involving interviewing, an estimate of interviewer variation should be done to help assess the reliability of the data collected. Such analysis will not only help judge interviewer performance, but in some cases will point to parts of the interview that need to be improved to obtain more reliable information.

TABLE 14.—*The interviewer variance for the three procedures during both time periods as measured by ρ which is the coefficient of intraclass correlation*

In the analysis of variance table:
the mean square for "among interviewers" $= s_b^2 + ks_a^2$
the mean square for "within interviewers" $= s_b^2$

$$\rho = \frac{s_a^2}{s_a^2 + s_b^2}$$

Period	Procedure	ρ
1	A	0.218
	B	.287
2	A	.091
	B	−.011
	C	−.023

With 4 and 30 degrees of freedom and k=8:
p=.05 when ρ=.175
p=.1 when ρ=.125

One other area in relation to interviewing is difficult to assess: the integrity of the interviewers. The variation in the total number of persons under study, that you may have noticed in the various tables presented, was due to one interviewer failing to complete her assigned interviews. Such an incident is unfortunate and probably rare, but it points out the need to evaluate interviewer performance.

The analyses presented here are some of the approaches that can be used to evaluate an interviewing procedure. Surveys for different purposes require different approaches. For example, if independent records of the events of interest are available, these will be valuable in evaluating the interview results. Regardless of the purpose for which an interview procedure is used, the performance of the procedure should be evaluated and not taken at face value.

Interviewing people is an instrument or tool that we use to obtain information. Like any tool we should know how to use it and what to use it for. A close analogy can be drawn between interviewing and another type of tool that we commonly use, a laboratory test. Before instituting a laboratory test, such as serum cholesterol or blood glucose determinations, we first must determine the best specific method to use. Usually there are methods described in the literature, and sometimes these have even been evaluated. On occasion, we may have to develop a new method. In either case, we must determine how the test performs in our own laboratories, that is, how accurate and reproducible it is in our hands, and what variation we can expect. Once the test is established, we want to be certain that our technicians are doing a good job. Although we don't often admit it, it is a good idea to check out our technicians once in a while to be sure that the test results are as good as we think they are. When this type of information is known, we can use the test and make meaningful statements about the results obtained.

The interviewing procedures we use should be evaluated with the same care as our laboratory tests. Furthermore, when a study that has used a laboratory test is published, the methods used are described so that the reader can better evaluate the work, and, if desired, try to repeat it. Similar information should be published about the interviewing techniques used in any study.

In summary, a small pilot study was presented in which three procedures for obtaining morbidity data from households have been compared. Two of these procedures were interviews; the third was a daily record. Each procedure was randomly assigned to one group of 50 families. All comparisons presented were based upon the acute illnesses having their onsets within the 4-week periods of observation. The highest rate of illness was reported on the daily records. More illnesses were reported in the interview that used a list of symptoms and diseases than in the interview that used questions about changes in activities, but the difference was of questionable significance. The rates for illnesses causing some change in activity were about the same for all three procedures. The problems presented by the use of a household respondent, the decline in reporting of illnesses with the passage of time, and the variation in results obtained by the interviewers were discussed.

REFERENCES

1. Allen, G. I., Breslow, L., Weissman, A., and Nisselson, H. 1954. Interviewing versus diary-keeping in eliciting information in a morbidity survey. *American Journal of Public Health. 44:* 919–927.
2. Cannell, C. 1961. Reporting of hospitalization in the health interview survey. Health statistics from the U.S. National Health Survey. United States Public Health Service Publication No. 584–D4.

3. Cannell, C. 1963. Comparison of hospitalization reporting in three survey procedures. Health Statistics from the U.S. National Health Survey. Public Health Service Publication No. 548–D8, U.S. Department of Health, Education, and Welfare.
4. Cartwright, A. 1959. Some problems in the collection and analysis of morbidity data obtained from sample surveys. *The Milbank Memorial Fund Quarterly. 37:* 33–48.
5. Downes, J., and Mertz, J. C. 1953. Effect of frequency of family visiting upon the reporting of minor illnesses. *The Milbank Memorial Fund Quarterly. 31:* 372–390.
6. Feldman, J. J. 1960. The household interview survey as a technique for the collection of morbidity data. *Journal of Chronic Diseases. 11:* 535–557.
7. Nisselson, H., and Woolsey, T. 1959. Some problems of the household interview design for the U.S. National Health Survey. *Journal of the American Statistical Association. 54:* 69–87.
8. Siegel, S. 1956. Nonparametric statistics for the behavioral sciences. McGraw-Hill Book Co., Inc., New York.
9. Snedecor, G. W. 1956. Statistical methods. Iowa State University Press, Ames, Iowa.
10. Velz, C. J., and Hemphill, F. M. 1953. Home injuries: investigation and application of home injury survey data in development of prevention procedures. University of Michigan School of Public Health, Ann Arbor, Mich.

PART IV

Current Epidemiological Problems

Intrafamilial Transmission of Rheumatoid Arthritis

By Sidney Cobb, M.D., *Survey Research Center, University of Michigan, Ann Arbor, Mich.*

"Perhaps the emphasis should be directed to environmental factors and to 'hereditary habits' rather than to hereditary tendencies" (Smith, 1928).

The above quotation comes from a paper on peptic ulcer, but it seems possible that it is equally applicable to rheumatoid arthritis. The literature on the familial aggregation of this disease has been adequately covered elsewhere (Edström, 1941; Stecher, 1957; McKusick, 1959; Blumberg, 1960; De Blecourt *et al.*, 1961; Lawrence, 1962; and Burch, 1963). Almost all studies bring out evidence for familial aggregation and it is usually assumed that the cause of this is genetic rather than environmental. At the present time the evidence is inadequate to support any simple genetic hypothesis. However, the genetic notion about this disease has persisted for about 100 years; therefore it is likely that there will be as much difficulty in eradicating it as in confirming it. What is needed is a strong alternative hypothesis for comparison with this genetic hypothesis.

The best alternative appears to be the psychosocial hypothesis. First, it is obvious that many things, in addition to genes, are passed on from father to son. Included among these things are property, religious preference and political affiliation. Of particular present relevance is the fact that personality characteristics and neurotic response patterns may be thus transmitted (Thorne, 1957; Ehrenwald, 1963). The reason for this relevance is the accumulating evidence that certain personality characteristics are fairly regularly associated with rheumatoid arthritis (King, 1955; Cobb, 1959; Moos, 1964). If, as has been suggested (Cobb, 1963a), the probability of this disease becoming manifest is increased by the interaction of these specific personality characteristics with relevant environmental situations,

then the rather small degree of familial clustering might possibly be explained without any genetic influence.

Briefly, incompletely and somewhat imprecisely stated, the theory of the psychosocial contribution to rheumatoid arthritis runs as follows. A sensitive, lively, responsive youngster, who is raised by a family in which the parent of the same sex is intolerant of aggression and excessively and unfairly punitive or critical, may learn to hide hurt feelings, bury aggressive responses and may fail to develop a stable self-evaluation. When later in life such an individual is put into a strong conflict situation that arouses anger and threatens self-esteem, the likelihood of active arthritis is increased. Other lively children of the same family may become angry, aggressive and rebellious. When later in life such persons are forced, by the power structure of the social environment in which they find themselves, to curb their anger and its expression, they are likewise more vulnerable to the disease. Thus we see that two kinds of personalities can be generated by the same parents and that both of these personalities may, under appropriate, but different circumstances, fill up with the negative self-evaluations and resentment which seem to contribute to rheumatoid arthritis. Since accumulation of resentment can lead to explosive and unexpectedly punitive behavior, the children of the person in question can be exposed to the same psychological environment. Thus there might be a propagation from generation to generation giving a semblance of genetic transmission. Though these attitudes and behaviors have been demonstrated to be related to rheumatoid arthritis, as mentioned above, an adequate connection to the disease via physiological mechanisms has not been established. We do, however, know enough about the physiological alterations associated with emotional states to make a fairly strong assumption that relevant physiologic processes exist. It is therefore an important research task to seek out the relevant mechanisms, for we will never be satisfied with a causal chain until all the links have been identified.

The present purpose is to see if we can explain the family clustering, not the entire causation of the disease. Our approach to this explanation is to select certain areas for study where the deductions from the two hypotheses, genetic and psychosocial, might give sharply contrasting results. So far six such areas have been identified. They are laid out in table 1. The first two areas are ones with regard to which there is already some information. First, Burch (1963), has found little, if any, evidence of familial clustering in a tribe of American Indians in which the frequency of the disease is high. Second, the implication of recent studies of an employed population under continuous observation for 28 months is that probably everyone is susceptible to this disease and in fact will probably have at least a brief attack if he lives long enough. Furthermore, it is noted under No. 6 that husbands and their wives are both affected more commonly than one would expect by

chance. Each of these points is consistent with the psychosocial hypothesis, though they by no means prove it. Of the remaining areas we will focus only on No. 6 which is the principle subject of a current investigation. Incidentally information on Nos. 4, 5, and 6 will be obtained also.

Table 1.—*A comparison of two hypotheses about the intrafamilial transmission of rheumatoid arthritis*

Genetic Hypothesis	Psychosocial Hypothesis	Facts
1. Family clustering should be found in all cultures.	In those cultures in which the sources of resentment and depression lie primarily outside the family there should be little familial clustering.	Little family clustering in American Indians (Burch, 1963).
2. Susceptibility should be limited by gene frequency.	Susceptibility should be unlimited.	Susceptibility appears unlimited (Lincoln & Cobb, 1963).
3. Given a propositus with RA the probability that a given relative will have RA should be dependent on the degree of consanguinity.	Given a prospositus with RA the probability that a given relative will have RA should be dependent on his psychological and social characteristics.	
4. Given two siblings, one with RA and the other without, the choice of which is affected should be independent of the characteristics of the spouse and the marital interaction pattern.	Given two siblings, one with RA and the other without, the choice of which is affected should be related to the characteristics of the spouse and of the marital interaction pattern.	Swaim (1962) has strong clinical impressions that the marital interaction pattern is important.
5. Given a propositus with RA, the probability of a sib having RA should be more dependent on the RA status of the parents than on any other characteristic of the parents.	Given a propositus with RA the probability of a sib having RA should be more dependent on the punitive characteristics than on the RA characteristics of the parents.	
6. Given a propositus with RA any excess frequency of RA in the spouse should be attributable to a mate selection effect and should show up in a sibling's spouse.	Given a propositus with RA the probability that the spouse will be affected should be greater than the general expectation and be dependent on the nature of the marital interaction pattern.	Spouses are affected more commonly than one would expect by chance (Holsti & Rantasalo, 1936; Schmid & Slatis, 1962; Cobb, 1963b).

Before going on to this matter it is appropriate to review our present picture of rheumatoid disease. There seem to be eight relevant points.

1) This is a systemic disease that may affect essentially every organ in the body. It is commonly called rheumatoid arthritis because most of the complaints are referred to the joints. Fortunately for our purposes, persons who have the disease without joint symptoms are probably rare.

2) It has a continuous and very gradual severity gradient. In the first place, it seems likely that everyone is susceptible though perhaps not equally so (Lincoln and Cobb, 1963). In the second place, it is very common and only a small percentage of those who have episodes of the disease are significantly disabled.

3) It is basically episodic in nature. Only the severe cases are continuously affected for long periods of time. This makes for real difficulty in diagnosis because a single examination will reveal only a small proportion of those who are intermittently affected.

4) It increases in frequency with age.

5) There is no difference in frequency in severe or in total disease between the sexes. However, moderate joint complaints are more common in middle-aged women.

6) The disease is more prevalent in spring and fall than at other seasons of the year.

7) The etiology is unknown and is probably multifactorial.

8) This is a disease that people like to talk about for there appears to be no social stigma attached to it. In fact many of its victims appear to be benefited by talking to a sympathetic and patient person. On the other hand, authoritarian individuals find victims uncooperative. This tells us something about the selection and training of interviewers for the study which we are proposing.

THE DESIGN

The design of the study involves people having three degrees of consanguinity and four kinds of relationships. Specifically the characteristics of the spouse, a sibling, a cousin and an individual unrelated to the propositus will be studied. Since the characteristics include the marital situation, it will be necessary to interview the spouse of each. Since the interview method will be used, the sample will be limited to people who can speak and read the English language. Furthermore, since the sample will probably not be large enough to analyse Negroes separately, only Caucasians will be included. The final sample will probably be composed of equal numbers of propositi selected a) from teaching hospital arthritis clinics, b) as severe cases in screening from national samples, and c) as mild cases from the national screening operation. The propositi from the national samples will be selected on the basis of the index of rheumatoid arthritis and three additional questions pertaining to diagnosis, severity and duration.

The relevant questions will be found in appendix A which is the screening instrument used in the first two national sample surveys. These surveys are ones that are conducted by the Survey Research Center for a variety of purposes at least four times a year and it is possible to add a few questions to each of these at moderate cost, where doing national screening for the purposes of this study alone would be prohibitively expensive. Each of these surveys is of a random sample of the continental United States and is drawn on the basis of a stratified area sampling technique (Kish and Hess, 1959).

It is appropriate at this point to ask about the validity of the index of rheumatoid arthritis as a screening instrument. As mentioned above, this index is composed of the first three questions in appendix A. It has been independently validated in two earlier studies (Rubin *et al.*, 1956; Cobb *et al.*, 1959) and it has been demonstrated that its sensitivity for probable plus definite rheumatoid arthritis lies between 66 percent and 77 percent, and its specificity between 95 percent and 98 percent. These screening characteristics are at least as good as those of the best screening tests for diabetes (Diabetes Pro-

gram Guide, 1960). In order to compensate for the slight lack of specificity, that is to eliminate the false positives, the second three questions in appendix A were added. These make it possible to exclude persons with other forms of arthritis, such as gout, and persons whose apparent joint swelling is due to peripheral edema. In addition they allow us to eliminate cases who have had but a single joint swollen or who have had only brief episodes of joint swelling. At subsequent interviews greater detail will be obtained in order to get at further reasons for exclusion and to measure degree of activity and severity.

Actually there will be three interviews at 4-month intervals. This will eliminate the seasonal variation and make it possible to estimate proportion of time in episode of active disease according to the method of Beall and Cobb (1961), and to collect the necessary data on the psychological and social factors.

Each family cluster will be made up of the 1) propositus with rheumatoid arthritis, 2) his spouse, 3) a sibling, 4) a maternal cousin, 5) a paternal cousin and 6) an unrelated individual. Individuals Nos. 3–6 will be selected for being within 5 years of age of the spouse of the propositus so that there will be no more than 10 years difference in ages between the cluster members and the difference in age between the propositus and the cluster members will be approximately constant. We have found that it is not possible to select any significant number of clusters in which all the cluster members are of the opposite sex from the propositus which would eliminate the need for adjustment for differences between the sexes, or separate analyses by sex. In all probability the more appropriate of these two approaches will be an adjustment for sex because, as noted above, it is believed that the fundamental frequency of rheumatoid disease is about the same in the two sexes but the arthritic manifestations are more prominent in middle-aged females.

In an ideal cluster all of the individuals will be married and living with their spouses and clusters will not be accepted unless the propositus, the sibling, the unrelated individual, and at least one of the cousins is living with the spouse. Present intentions are not to accept any propositi under 35 and it seems unlikely that there will be many propositi over 65 because most persons of this age will have lost one or more relevant persons from the cluster by death. Obviously, the older clusters will be the most desirable for in them the penetrance will be most nearly complete. Very preliminary information from the first screening operation suggests that we will be able to locate slightly over 10 ideal clusters per 1,000 adults interviewed. The procedures for sampling patients under treatment have not yet been worked out but they should be relatively simple and will be focused on members of the American Rheumatism Association who have teaching hospital appointments. Optimum sample size

has not yet been determined, but it is estimated that it will be of the order of 400 clusters.

MEASUREMENT

In order to determine which of the proposed hypotheses is more correct it will be necessary to measure each member of each cluster on three dimensions: the degree of consanguinity, the degree of rheumatoid arthritis, and the psychological and social characteristics. We can quickly dispose of the problem of measuring consanguinity by realizing that siblings on the average share ½ of their genes while cousins share only ⅛ of their genes, and presumably the unrelated individuals share only a negligible proportion of their genes with the propositus. This gives us three degrees of consanguinity and if by chance we should find any instances of double cousins a fourth degree will be available. For present purposes it is appropriate to assume that the errors of accepting the apparent paternity are unimportant in comparison to the measurement errors that will appear in the other variables. The instrument used for the identification of the members of the cluster is presented as appendix B.

The measurement of rheumatoid arthritis will be divided in three parts. The first is the determination of the diagnosis. Here table 2 is relevant for it shows the relationship of various measures to the most accurate diagnosis possible on persons under continuous observation for 28 months. This table speaks for itself and demonstrates clearly that the rheumatoid arthritis index used on three separate occasions is a better measure of rheumatoid arthritis than any serological test or single clinical examination. The reasons for this are that rheumatoid arthritis is a remittant disease which is by no means all discovered at a single examination, and the best serological tests are as yet insufficiently sensitive to be really useful in a study of this sort. The specificity of the index will be increased by using a set of questions designed to exclude other rheumatic diseases and joint swelling due to peripheral edema. This should give specificity ap-

TABLE 2.—*The sensitivity and specificity of various measures of rheumatoid arthritis for the best clinical diagnosis that can be made after monthly screening for 28 months*

Measure	For definite RA		For probable and definite RA	
	Sensitivity	Specificity	Sensitivity	Specificity
Serology IRLF*	28	91	14	92
Serology PLAT*	50	90	30	93
Single Exam Prob. & Def.	13	100	16	100
Single Exam Poss. & Prob. & Def.	69	89	51	92
RA index positive at least one out of three times	75	91	53	95

*Brooks and Cobb, 1963.

proaching that of the best clinical evaluations which are also subject to some false positives, as pointed out by Dixon (1960).

The second part of the measurement of the arthritis is the determination of the degree of activity. This will be determined from four things: the number of joints swollen, the duration of morning stiffness, the amount of aspirin and related drugs consumed, and the proportion of time in episode with active disease. The third part is the measurement of the degree of disability resulting from the arthritis. This will be estimated from a series of questions about specific musculo-skeletal functions, as well as questions about time off duty, time in the hospital, frequency of visits to physicians or clinics and need for special treatments or appliances. Finally these three measures will be put together with appropriate weights to give an RA score which will hopefully have quite a wide range and considerable validity. Its validity will be checked on the propositi sampled from persons under treatment for this disease.

It is not within the scope of this paper to discuss the psychosocial variables in detail, however, it is appropriate to point out that the general characteristics outlined in the psychosocial hypothesis will be studied in five categories: demographic characteristics, preception of parents, personal characteristics, marital interaction and general environmental frustration. Insofar as possible, well-established psychological measures will be used, but for certain items it will be necessary to design appropriate measurement techniques of our own. The reliability of measures of this sort is for the most part very high. The difficulties lie in the area of validity. Take for example the index of impulsiveness versus control, which has been shown to be significantly related to rheumatoid arthritis (Cobb *et al.*, 1963). To begin with, the concept of impulsiveness is ill-defined and there is no standard measure of this. On the other hand, the questions used all have good face validity in that they are direct and apparently relevant, as the true-false question "I often act on the spur of the moment without stopping to think"; and they are all modestly intercorrelated. This means that they can be used to make a scale which is reasonably satisfactory though it may not have equal intervals (Likert, 1932). Since there is now a fairly substantial literature on the psychological and social factors in rheumatoid arthritis, it seems likely that sound and appropriate measures can be worked out that will cover the relevant variables sufficiently well to give significant results. It is useful to remember that in every field of measurement we inevitably fall somewhat short of "God's opinion." The important thing is to determine if a given measurement is sufficiently valid and reproducible to contribute to scientific progress.

The interviews will be conducted by the regular members of the Survey Research Center's nationwide interview staff, using carefully worked out and pretested forms. Since interviewer bias could give a

semblance of familial clustering if each cluster were to be assigned to a single interviewer, an attempt will be made to distribute the interviews in a cluster among several interviewers. Coding will be done in accordance with the usual standards of the center (Muehl, 1961). Transfer to IBM punch cards will make possible the complex analyses to follow.

ANALYSIS

In table 3 the three main lines of analysis are laid out. The first analytical job will be to put together the RA score. Then we will assemble the various psychological and social measurements into an index indicative of the psychosocial hypotheses stated above. This will be done *a priori* on the basis of knowledge from previous studies rather than on the basis of their value in identifying persons with rheumatoid arthritis in this study. This will avoid the possibility of developing a spuriously strong relationship to these items. It will be necessary to work out a procedure for adjusting for the differences in reported arthritis between the sexes. This should not be difficult if a sound decision can be arrived at as to whether to use an additive or a multiplicative model. In addition it will be necessary to correct for ascertainment (Crow, 1963) and check for bias in the selection of cousins. Procedures for these purposes are being worked out.

TABLE 3.—*The analysis*

A. Mean RA score by consanguinity and psychosocial score:

Relationship	Degree-of-Psychosocial Characteristics			Overall
	High	Medium	Low	
Sibs				
Cousins				
Isolates				
Overall				

B. Analysis of variance in the same cells.
C. Multiple regression analysis using the following continuous variables:
   a. RA score.
   b. Degree of consanguinity.
   c. Psychosocial score.

The first analysis will be to examine the mean arthritis score for the several cells in table 3. Then an analysis of variance of the rheumatoid arthritis score will be undertaken for the same table. It is recognized, of course, that this is fraught with the usual dangers of doing an analysis of variance in a situation in which the assignment to the various cells has not been made at random. However, the results are likely to prove instructive. Finally a multiple regression analysis will be undertaken using the three continuous variables: rheumatoid arthritis score, degree of consanguinity, and the psychosocial score. Separate analyses will be done by source of case, whether from the patient sample, from the national sample of severe cases or the national

sample of mild cases. Then all types of cluster will be analyzed together.

DISCUSSION

Continuing to focus only on hypothesis No. 3 in table 1, it is perfectly clear that there is a substantial chance of an indeterminate answer when in fact the true answer should be primarily genetic or primarily psychosocial. The likelihood of this is reduced by increasing the validity of our measurements and by increasing the sample size (Rubin *et al.*, 1956). Even with the best measures and optimum sample size it would be unrealistic to expect an answer on this deduction which says more than "the facts seem to be more in line with the genetic hypothesis or more in line with the psychosocial hypothesis." At this point examination of all the deductions will play a useful role for if the results are consistent across a whole series of deductions the conclusion will be much firmer. The results of this study will not absolutely exclude either hypothesis, though they might suggest that the contribution of one or the other is sufficiently small to be disregarded. Of course, it is possible that both hypotheses are correct and each accounts for a substantial portion of the variance, in which case the proportion of each will be estimated.

This design gets us out of the problem of studying all the members of a large kinship which in fact may be thoroughly atypical. By contrast it puts us in a situation in which we get bits of data from a wide distribution by social class, subculture and geographic area. This means that any conclusions that are drawn from this kind of a study can much more likely be generalized to the population than are those derived from intensive studies of single kinships. Surely this general kind of design is applicable to a wide variety of genetic and other intrafamilial problems.

SUMMARY

Rheumatoid arthritis appears to cluster in families. This may be due to genetic or interpersonal factors or in part to both. A design for examining this problem has been presented. This design is applicable to a wide variety of similar problems.

ACKNOWLEDGMENTS

The assistance of Ernest Harburg, who managed the initial field work of this study, and of W. J. Schull, who is consultant to the study, is gratefully acknowledged.

The research here reported has been supported in part by the following grants from the National Institutes of Health: AM–06928, MH–K6–16, 709, MH–09177.

735–712 O—65——15

AGE_____ SEX_____ RACE_____ COOPERATION_____ RATING_____

	COL. A	*COL. B*
R1. Have you ever at any time had arthritis or rheumatism?	☐ YES (OR DON'T KNOW)	☐ NO
R2. Do you wake up with stiffness or aching in your joints or muscles?	☐ YES, USUALLY OR SOMETIMES	☐ NO (OR DON'T KNOW)
R3. Have you ever had any swelling in any joints?	☐ YES	☐ NO (OR DON'T KNOW)

R4. (INTERVIEWER—CHECK ONE)

☐ *ALL* BOXES IN COLUMN A CHECKED (GO ON WITH Q. R5)
☐ *ONE* OR MORE BOXES IN COLUMN B CHECKED—(SKIP TO PAGE 47, Q. T1—DO *NOT* LEAVE ARTHRITIS LETTER OR FORM I)

R5. Do you think this swelling was due to arthritis? _____

_____

R5a. Why is that? _____

_____

R6. What joints are or were swollen? Please tell me which ones. _____

_____

R6a. Any others? _____

R7. What was the longest period of time when your joints were continuously swollen? _____

R8. (INTERVIEWER—CHECK ONE)

MY RESPONDENT IS MARRIED AND LIVING WITH SPOUSE (from Q. P2, Page 39)
☐ YES—(GO ON WITH Q. R9)
☐ NO—(SKIP TO PAGE 47, Q. T1—DO *NOT* LEAVE ARTHRITIS LETTER OR FORM I)

R9. Next, we'd like to ask you, do you have any married brothers (*NOT* step-brothers or half-brothers)?
☐ YES ☐ NO—(GO ON WITH PAGE 45, Q. R10)

R9a. What are the ages of those brothers who are now living with their wives?
_____ _____ _____ _____ _____ _____

R10. Do you have any married sisters (*NOT* step-sisters or half-sisters)?
☐ YES ☐ NO—(GO TO Q. R11)

R10a. What are the ages of those sisters who are now living with their husbands?
_____ _____ _____ _____ _____ _____

R11. (INTERVIEWER—CHECK ONE)

☐ R HAS *NO* BROTHERS *OR* SISTERS WHO *ARE MARRIED AND LIVING WITH SPOUSE*—(SKIP TO PAGE 47, Q. T1—DO *NOT* LEAVE ARTHRITIS LETTER OR FORM I)
☐ R HAS A BROTHER OR A SISTER (OR BOTH) WHO IS MARRIED AND LIVING WITH SPOUSE—(GO ON WITH Q. R12)

R12. Do you have any first cousins?
☐ YES—(CONTINUE WITH Q. R13) ☐ NO—(GO TO Q. R14, BELOW)

R13. Are these first cousins about your age, that is, within ten years of your age?
☐ YES ☐ NO—(GO TO Q. R14, BELOW)

R13a. How many first cousins about your age are there on your *mother's* side?
☐ONE ☐TWO ☐THREE ☐FOUR ☐FIVE OR MORE
R13b. How many first cousins about your age are there on your *father's* side?
☐ONE ☐TWO ☐THREE ☐FOUR ☐FIVE OR MORE
(GO ON WITH Q. R14)

THE UNIVERSITY OF MICHIGAN
INSTITUTE FOR SOCIAL RESEARCH
ANN ARBOR, MICHIGAN
RENSIS LIKERT, DIRECTOR

Survey Research Center
Angus Campbell, Director
Research Center for Group Dynamics
Alvin Zander, Director

MAY 1963.

A LETTER TO PERSONS WITH ARTHRITIS:

One aim of the interview you have just completed is to find people with arthritis. Even if your arthritis is very mild, you can help a great deal in this nationwide survey.

This survey is to find some of the causes of arthritis. Why is arthritis more common in some families than in others? To aid the millions who bear the pain of this illness, we must try to find answers to this problem.

You can help in this search by telling us how to get in touch with other members of your family. We need to get the names and addresses of your brothers and sisters and some of your cousins. We will then select a few of them and talk with them about their health.

Not many of us carry these addresses around in our heads, so I have asked your interviewer to leave the attached form with you. In this way you will have time to fill in the names and addresses for the study. Let me assure you that all the information you give us will be held *confidential.*

Your help will be appreciated by many people—those who suffer from arthritis as well as their friends and families.

Sincerely yours,

SIDNEY COBB, M.D.,
Program Director.

FORM I Page 1

1. Please print your name here _____ (Give middle initial)

2. First, we would like to ask you about your BROTHERS and SISTERS. (*NOT* your step-brothers or step-sisters. *NOT* your half-brothers or half-sisters.)

   Simply write in the full name, age and present address of each of your brothers and sisters who are now living.

3. START HERE:

(1) Please write the full name of each brother or sister, one in each box. (Give married name)	(2) How old is he, or she	(3) What is his or her present address? STREET AND NUMBER CITY AND STATE PHONE NUMBER	(4) Is he or she married? (Check one)
First name Last name	Age	PHONE NUMBER:	☐ Married, lives with spouse ☐ Single ☐ Widowed ☐ Divorced or separated
First name Last name	Age	PHONE NUMBER:	☐ Married, lives with spouse ☐ Single ☐ Widowed ☐ Divorced or separated
First name Last name	Age	PHONE NUMBER:	☐ Married, lives with spouse ☐ Single ☐ Widowed ☐ Divorced or separated
First name Last name	Age	PHONE NUMBER:	☐ Married, lives with spouse ☐ Single ☐ Widowed ☐ Divorced or separated

(LIST ANY OTHER BROTHERS OR SISTERS ON BACK OF THIS PAGE)

6–63/RA/303 PLEASE GO ON WITH NEXT PAGE

1. Now think about your *father's brothers.* ☐ (IF NONE, CHECK BOX AND GO TO *PART B* BELOW.)
2. Think about the *children* each of your father's brothers may have had. These are your first cousins. ☐ (IF NONE, CHECK BOX AND GO TO *PART B* BELOW.)
3. Choose *one* of these cousins who is *closest in age to your wife or husband.*
   a. My wife/husband is _____ years of age.
   b. The cousin I have in mind who is closest to this age is _____ years of age, and is (check one) ☐ MALE ☐ FEMALE
   c. His/her full (married) name is: _____
   d. You can write to him/her at this address:
      NUMBER AND STREET: _____
      CITY, ZONE, AND STATE: _____
   e. The *phone number* is: _____ (GO ON WITH PART B)

PART B

1. Now think about your *father's sisters.* ☐ (IF NONE, CHECK BOX AND GO ON WITH NEXT PAGE.)
2. Think about the *children* each of your father's sisters may have had. These are your first cousins. ☐ (IF NONE, CHECK BOX AND GO ON WITH NEXT PAGE.)
3. Choose *one* of these cousins who is *closest in age to your wife or husband.*
   a. My wife/husband is _____ years of age.
   b. The cousin I have in mind who is closest to this age is _____ years of age, and is (check one) ☐ MALE ☐ FEMALE
   c. His/her full (married) name is: _____
   d. You can write to him/her at this address:
      NUMBER AND STREET: _____
      CITY, ZONE, AND STATE: _____
   e. The *phone number* is: _____

PART C

1. Now think about your *mother's brothers.* ☐ (IF NONE, CHECK BOX AND GO TO *PART D* BELOW.)
2. Think about the *children* each of your mother's brothers may have had. These are your first cousins.
   ☐ (IF NONE, CHECK BOX AND GO TO *PART D* BELOW.)
3. Choose *one* of these cousins who is *closest in age to your wife or husband.*
   a. My wife/husband is_____years of age.
   b. The cousin I have in mind who is closest to this age is_____years of age, and is (check one): ☐ MALE ☐ FEMALE
   c. His/her full (married) name is:_____
   d. You can write to him/her at this address:
      NUMBER AND STREET:_____
      CITY, ZONE, AND STATE:_____
   e. The *phone number* is:_____ (GO ON WITH PART D)

PART D

1. Now think about your *mother's sisters.* ☐ (IF NONE, CHECK BOX AND GO ON TO NEXT PAGE.)
2. Think about the *children* each of your mother's sisters may have had. These are your first cousins.
   ☐ (IF NONE, CHECK BOX AND GO ON WITH NEXT PAGE.)
3. Choose *one* of these cousins who is *closest in age to your wife or husband.*
   a. My wife/husband is_____years of age.
   b. The cousin I have in mind who is closest to this age is_____years of age, and is (check one): ☐ MALE ☐ FEMALE
   c. His/her full (married) name is:_____
   d. You can write to him/her at this address:
      NUMBER AND STREET:_____
      CITY, ZONE, AND STATE:_____
   e. The *phone number* is:_____

In each family there is usually a relative who is most likely to know the names and addresses of most of the members of the family.

PART E

1. The person in my family who would most likely know the names and addresses of most of my FATHER'S family is:

FULL NAME: _____

STREET ADDRESS: _____

CITY or TOWN: _____

STATE: _____

PHONE NUMBER: _____

2. The person in my family who would most likely know the names and addresses of most of the members in my MOTHER'S family is:

FULL NAME: _____

STREET ADDRESS: _____

CITY or TOWN: _____

STATE: _____

PHONE NUMBER: _____

Thank you for your kind cooperation. We will write to you soon, telling you more about our study. In the meantime, any comments you may have will be sincerely appreciated.

_____

_____

_____

_____

SIDNEY COBB, M.D.,
Survey Research Center, The University of Michigan.

PLEASE KEEP THIS FORM. OUR INTERVIEWER WILL RETURN TO PICK IT UP IN THE NEAR FUTURE.

5–63/RA/P.303

REFERENCES

1. Beall, G. and Cobb, S. 1961. The frequency distribution of episodes of rheumatoid arthritis as shown by periodic examination. J. Chron. Dis. 14: 291–310.
2. De Blecourt, J. J., Polman, A. and De Blecourt-Meindersma, T. 1961. Hereditary factors in rheumatoid arthritis and ankylosing spondylitis. Ann. Rheum. Dis. 20: 215–223.
3. Blumberg, S. 1960. Genetics and rheumatoid arthritis. Arth. & Rheum. 3: 178–185.
4. Brooks, G. W. and Cobb, S. 1963. Studies in latex agglutination: An approach to the determination of optimum conditions for discrimination between rheumatoids and normals. Arth. & Rheum. 6: 198–207.
5. Burch, T. 1963. (This symposium.)
6. Cobb, S. 1959. Contained hostility in rheumatoid arthritis. Arth. & Rheum. 2: 419–425.
7. Cobb, S. 1963a. The epidemiology of rheumatoid arthritis. Bull. N.J. Acad. Med. 9: 52–60.
8. Cobb, S. 1963b. Unpublished data from the Pittsburgh Arthritis Study.
9. Cobb, S., Chen, E. and Christenfeld, R. 1963. Psychological and social characteristics of persons with various psychosomatic diseases. (Read before Ann. Meeting Am. Psychosom. Soc., April 1963.)
10. Cobb, S., Miller, M. and Wieland, M. 1959. On the relationship between divorce and rheumatoid arthritis. Arth. & Rheum. 2: 414–418.
11. Crow, J. F. (This symposium.)
12. Dixon, A. St. J. 1960. "Rheumatoid arthritis" with negative serological reaction. Ann. Rheum. Dis. 19: 209.
13. Edström, G. 1941. Klinische Studien über den chronischen Gelenkrheumatismus. Acta Medica Scand. 108: 398–413.
14. Ehrenwald, J. 1963. *Neurosis in the family and patterns of psychosocial defense. A study of psychiatric epidemiology.* Hoeber, New York.
15. Holsti, O. and Rantasalo, V. 1936. On the occurrence of arthritis in Finland. Acta Med. Scand. 88: 180–195.
16. King, S. H. 1955. Psychosocial factors associated with rheumatoid arthritis: An evaluation of the literature. J. Chron. Dis. 2: 287–302.
17. Kish, L. and Hess, I. 1959. The SRC national sample of dwellings. Survey Research Center, Mimeograph.
18. Lawrence, J. S. 1962. Genetic studies on rheumatoid arthritis. A.J.P.H. 52: 1689–1696.
19. Likert, R. 1932. A technique for the measurement of attitudes. Arch. Psychol. No. 140.
20. Lincoln, T. A. and Cobb, S. 1963. The prevalence of mild rheumatoid arthritis in industry. J. Occ. Med. 5: 10–16.
21. McKusick, V. A. 1959. Genetic factors in diseases of connective tissue. Am. J. of Med. 26: 283–302.
22. Muehl, D. 1961. A manual for coders. Survey Research Center, Ann Arbor, Michigan.
23. Moos, R. H. 1964. Personality factors associated with rheumatoid arthritis: A review. J. Chron. Dis. 17: 41–55.
24. Rubin, T., Rosenbaum, J. and Cobb, S. 1956. The use of interview data for the detection of associations in field studies. J. Chron. Dis. 4: 253–266.
25. Schmid, F. R. and Slatis, H. 1962. Increased incidence of serum and clinical abnormalities in spouses of patients with rheumatoid arthritis (Abstract). Arth. & Rheum. 5: 319–320.

26. Smith, D. A. 1928. A statistical review of gastric and duodenal ulcer. Brit. Med. J. 2: 293–297.
27. Stecher, R. M. 1957. Heredity in joint disease. Documenta Rheumatologica, J. R. Geigy S.A., Basel, Switzerland.
28. Swaim, L. 1962. *Arthritis medicine and the spiritual laws.* Chilton, New York.
29. Thorne, F. C. 1957. Epidemiological studies of chronic frustration-hostility-aggressive states. Am. J. Psychiat. 113: 717–721.
30. U.S. Department of Health, Education, and Welfare. 1960. Diabetes Program Guide. PHS Publication No. 506. U.S. Government Print. Office, Washington.

Family and Genetic Studies of Rheumatoid Arthritis and Rheumatoid Factor in Blackfeet Indians

By Thomas A. Burch, M.D., M. P. H., William M. O'Brien, M.D., and Joseph J. Bunim, M.D., *National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare, Bethesda, Md.*

From 1958 to 1962 (Ziff *et al.*, 1958; Lawrence and Ball, 1958; Bremner *et al.*, 1959; Goldenberg *et al.*, 1960; and Schmid and Slatis, 1962) five family studies on the occurrence of rheumatoid factor (RF) and rheumatoid arthritis (RA) in first-degree relatives (parents, siblings and children) of patients with seropositive and seronegative RA have been reported in the English literature from this country, England and Scotland. Only Bremner *et al.* (1959) failed to find a significant aggregation of RA and/or RF in families of seropositive rheumatoid patients.

Unfortunately, important differences exist in the methods and material in these studies and in view of these fundamental differences, it is surprising that greater discrepancies in the results did not follow.

The present study was designed to answer the following questions concerning rheumatoid factor and rheumatoid arthritis.

1. *Is there a familial aggregation in a random sample of the selected population?* This will be analyzed by contrasting the frequency of RF and RA in first-degree relatives of probands with and without RF and RA, with the frequency in a random sample of the population tested.

2. *Is there any evidence of a simple genetic mechanism?* This will be analyzed by comparing the observed frequency of RF and RA in offspring of phenotypically different types of parents, and in sibling

pairs with the frequency that could be expected from the various genetic hypotheses and by chance alone.

The Blackfeet Indian population in Montana was especially suited for this type of survey for several reasons. Most of the relatives of any one individual lived in the same area; there was no reluctance to give truthful histories as to parent-child relationship and first cousin matings were very rare.

MATERIAL

In the fall of 1961 we conducted an area survey of Indians listed on the Blackfeet tribal roll, above age 29, residing on the Indian reservation in Montana. Each respondent was questioned carefully and examined by one of three physicians. X-ray films of hands and feet were made. Serum samples were collected and tested for RF by both the Bentonite flocculation test (BFT) (Bozicevich *et al.*, 1958) and the sheep cell agglutination test (SCAT) done by the method of Ball* (1950). From this survey appropriate probands were selected. To complete the family studies it was found advisable to return to the reservation several months later. The second session brought the completion rate of the population survey to 86 percent and the family study to 89 percent.

FAMILY STUDY

The probands and their first-degree relatives were classified into four groups as shown in table 1. Group A consisted of 22 probands with seropositive RA; group B had 34 individuals with a positive BFT but without RA; group C consisted of 20 Blackfeet Indians with seronegative RA and group D probands who were seronegative without RA; were matched with groups A and B probands by age and sex.

An additional control group was obtained by *randomly* matching each of the first-degree relatives of the positive probands (groups A, B and C) with someone from the general population survey of the same age and sex without regard to their clinical or serological status.

The prevalence of rheumatoid factor (tested by the BFT) and of rheumatoid arthritis in first-degree relatives of the four groups of probands and in the random sample are shown in table 1. The

*In addition to the studies described above we are currently completing analysis of results of the following procedures carried out in connection with this survey: 1) X-rays of the cervical spine, hips and sacro-iliac joints; 2) serum uric acid concentration; 3) major and minor blood group typing done by Dr. G. Albin Matson of the Minneapolis War Memorial Blood Bank; 4) ABO salivary secretor status done by Dr. E. J. Holborow of the Canadian Red Cross Memorial Hospital, Taplow, England; and 5) serum Gm groups determined by Dr. Paul Alepa in our laboratory.

differences between the various groups was no greater than would occur by chance ($p>.40$ and $>.90$ respectively).

When the SCAT was used instead of the BFT to test for RF the differences between the various groups were more marked but still were not significant.

TABLE 1.—*Prevalence of rheumatoid factor and rheumatoid arthritis in relatives of probands and in a random sample*

Probands				1° Relatives		
Group	BFT	RA*	Number	Number	Positive BFT	Rheumatoid arthritis*
A	positive	Yes	22	71	1 (1.4%)	2 (2.8%)
B	positive	No	34	77	4 (5.2%)	3 (3.9%)
C	negative	Yes	20	63	4 (6.3%)	2 (3.2%)
D	negative	No	55	179	5 (2.8%)	5 (2.8%)
Matched controls				211	12 (5.6%)	9 (4.3%)

*Probable and definite RA.

GENETIC ANALYSIS

In addition to the family study just described we were able to take advantage of the completeness of the survey and analyze the data from the numerous kinships on the reservation to see if they would fit a genetic hypothesis. In order to apply the Hardy-Weinberg Law to offspring of genotypically different parents, the mating in the population must be random in regard to the characteristic under study and there should be no selection factor that would limit the size of the families. Comparison of the observed numbers of matings with neither, one, or both members affected with the numbers to be expected on a random basis shows no significant difference for either RF or RA ($p>.95$ for both) (table 2). Furthermore there was but little difference in the number of offspring from affected and unaffected parents. In fact spouse pairs in which one member had RA had more offspring over the age of 29 (1.57 per pair) than those with both spouses negative (1.05 per pair).

TABLE 2.—*Occurrence of rheumatoid factor and rheumatoid arthritis in spouse of Blackfeet Indians*

	In both spouses	In one spouse	In neither spouse
Rheumatoid factor:			
Observed	2	32	218
Expected*	1.28	33.43	217.38
Rheumatoid arthritis:			
Observed	1	20	254
Expected**	0.44	21.12	253.44

*Based on prevalence of 7.1 percent positive BFT in 252 couples (504 persons) tested.
**Based on prevalence of 4.0 percent "probable" or "definite" RA in 275 couples (550 persons) examined.

The types of matings, frequency of matings and frequency of the various types of offspring in a large random mating population are shown in table 3. In this table A can be taken to represent the gene for RF or RA and a represents its allele but dominance is not necessarily implied. The gene frequency of A is p, and $1-p=q$, the gene frequency of a. Using the frequency of RF and of RA in the Blackfeet population, the values of p and q have been calculated according to the Hardy-Weinberg equilibrium both as if the trait were a dominant and as a recessive.

	Frequency	*Dominant*	*Recessive*
RF	0. 05882	$q^2=0.94118$	$p^2=0.05882$
		$q=0.97014$	$p=0.2425$
		$p=0.02986$	
RA	0. 03966	$q^2=0.96034$	$p^2=0.03966$
		$q=0.97996$	$p=0.19910$
		$p=0.02004$	

TABLE 3.—*Matings and offspring in a large random mating population*

Type of mating (reciprocals combined)	Frequency of mating	Offspring		
		AA	Aa	aa
AA × AA	p^4	p^4		
AA × Aa	$4p^3q$	$2p^3q$	$2p^3q$	
AA × aa	$2p^2q^2$		$2p^2q^2$	p^2q^2
Aa × Aa	$4p^2q^2$	p^2q^2	$2p^2q^2$	$2pq^3$
Aa × aa	$4pq^3$		$2pq^3$	q^4
aa × aa	q^4			
Total	1	p^2	2pq	q^2

Mating phenotypes	Ratio of offspring affected/total	
	Dominant	Recessive
A × A (pos. × pos.)	$\frac{3-2p}{(2-p)^2}$	1.00
A × a (pos. × neg.)	$\frac{1}{2-p}$	$\frac{p}{p+1}$
a × a (neg. × neg.)	0	$\frac{p^2}{(p+1)^2}$
A × ? (pos. × ?)	$\frac{1+p-p^2}{2-p}$	p
a × ? (neg. × ?)	p	$\frac{p^2}{1+p}$

The expected number of positive offspring for each mating class calculated both as a dominant and as a recessive trait are compared in table 4 with the observed numbers and the numbers expected if the distribution were entirely random. Chi-square tests show that only the random distribution fits the observed findings.

Study of the inheritance of rheumatoid arthritis is complicated by the late and variable age of onset of the disease. In analyzing the offspring of phenotypically different types of parents, as above, it should be borne in mind that some of the younger generation might

TABLE 4.—*Analysis of matings in Blackfeet Indians*

(Based on Hardy-Weinberg Law)

Parents		Offspring				
Type of mating	No.	Total examined	Observed positive	Expected number positive if—		
				Dominant* inheritance	Recessive* inheritance	Random distribution
Rheumatoid factor:						
Neg. × neg	102	112	4	0.000	2.603	5.419
Pos. × neg	12	15	1	4.547	1.786	0.725
Pos. × pos	2	1	0	0.453	0.610	0.048
Neg. × ?	127	288	16	7.644	12.119	13.935
Pos. × ?	9	18	0	8.356	3.880	0.870
X^2				∞	6.826	1.699
p				0	0.076	0.64
*Rheumatoid arthritis***						
Neg. × neg	112	128	4	0.000	1.965	3.413
Pos. × neg	14	22	0	4.000	2.034	0.586
Neg. × ?	130	294	10	2.960	7.730	10.723
Pos. × ?	13	35	2	9.040	4.260	1.276
X^2				∞	8.40	1.143
p				0	0.015	0.56

*The theoretical frequency is estimated from the assumed gene frequency which is based on observed prevalence of R.F. (5.882 percent) and R.A. (3.966 percent) in the 1961 survey.
**The theoretical frequency, adjusted proportionally to equal the number observed.

have been destined to develop RA but this, of course, would not be apparent at the time the data were collected. In other instances, by the time the children reached sufficient age to develop RA, the parents might be dead.

Obviously a method of analysis based on a single generation, who would be approximately the same age, would be a superior method of studying the inheritance of an age dependent trait. Such a method was devised by Penrose (1935) to study linkage and was extended by Cotterman (1937) to study unit factor inheritance. This is an extension of the Hardy-Weinberg law to cover the special case of sibling pairs. The frequency of each combination of sibling pairs for all possible types of matings is shown in table 5 with the trait considered in one instance dominant and in another recessive.

The number of pairs obtained from a family of size n is the number of combinations of n things taken two at a time or $n(n-1)/2$. Hence a family of 4 siblings would consist of $4\times3/2=6$ sibling pairs and 5 siblings would have $5\times4/2=10$ sibling pairs, etc. A total of 1,066 such sibling pairs were identified in the survey population. The number of pairs with neither, one, or both members affected is compared in table 6 with the number that would be expected assuming the trait were dominant, recessive or randomly distributed. As in the previous analysis the random distribution of affected individuals is the only one compatible with that actually observed.

TABLE 5.—*Matings and sibling pairs in a large random mating population*

Type of mating	Frequency of mating	Frequency of sibling pairs in offspring					
		Considering A as dominant			Considering A as recessive		
		RA-RA (S_{++})	RA-Norm (S_{+0})	Norm-Norm (S_{00})	RA-AR (S_{++})	RA-Norm (S_{+0})	Norm-Norm (S_{00})
AA × AA	p^4	p^4			p^4		
AA × Aa	$4p^3q$	$4p^3q$			p^3q	$2p^3q$	p^3q
AA × aa	$2p^2q^2$	$2p^2q^2$					$2p^2q^2$
Aa × Aa	$4p^2q^2$	$9/4p^2q^2$	$3/2p^2q^2$	$1/4p^2q^2$	$1/4p^2q^2$	$3/2p^2q^2$	$9/4p^2q^2$
Aa × aa	$4pq^3$	pq^3	$2pq^3$	$2pq^3$			$4pq^3$
aa × aa	q^4			q^4			q^4
Totals		$(1-S_{+0}-S_{00})$	$1/2pq^2(3+q)$	$1/4q^2(1+q)^2$	$1/4p^2(p+1)^2$	$1/2p^2(1-p)(3+p)$	$(1-S_{+0}-S_{++})$

TABLE 6.—*Analysis of sibling pairs in the Blackfeet Indians*

(Based on Hardy-Weinberg Law and Penrose Formula)

Type of pair	Observed	Expected number with—		
		Dominant* inheritance	Recessive* inheritance	Random distribution
Rheumatoid factor:				
Both negative	850	854.836	847.171	850.633
One positive	85	52.214	64.588	83.324
Both positive	1	28.950	21.241	2.040
	X^2	47.597	23.782	0.563
	p	<0.00001	<0.00001	0.45
Rheumatoid arthritis**				
Neither	983	973.562	996.644	981.442
One	81	59.466	54.160	82.828
Both	2	32.971	15.196	1.737
	X^2	36.980	24.936	0.081
	p	<0.00001	<0.00001	0.78

*The theoretical frequency is estimated from the assumed gene frequency which is based on observed prevalence of R.F. (5.882 percent) and R.A. (3.966 percent) in the 1961 survey.
**Probable or definite R.A. according to the A.R.A. criteria.

DISCUSSION

It is thus apparent that at least among the Blackfeet there is no evidence of familial aggregation of RF or RA and that the observed data are consistent with a randomly distributed condition and not consistent with simple dominant or recessive inheritance. A complex type of inheritance due to two recessive genes with a low degree of penetrance, has not been excluded but the evidence makes such types of inheritance very unlikely. Secondly, there could be a rare inherited form but, since there is no way of distinguishing this hypothetical inherited RA from the usual form of the disease, it is impossible to test this hypothesis.

It is difficult to reconcile the conflicting observations reported by other workers and those reported here but they may well be due to

differences in methods and material. In none of the published studies on familial aggregation, for instance, were the probands, relatives and controls all from exactly the same population. In only two studies (Stecher *et al.*, 1953, and Neri Serneri and Bartoli, 1956) was an attempt made to test genetic hypotheses and in both of these the method of ascertaining affected individuals among the relatives was inadequate.

We have recently completed a similar survey on the Pima Indians of southern Arizona. A preliminary perusal of the data indicates that they also do not have a familial aggregation of RA or RF. If the final analysis of these data is similar to that reported here for the Blackfeet, it will be necessary to search for a nongenetic etiology for rheumatoid arthritis and rheumatoid factor.

CONCLUSIONS

1. There is no evidence of familial aggregation of RF or RA in the Blackfeet Indians of Montana.
2. The frequency of offspring with RF and RA from "phenotypically" different types of parents and in sibling pairs was consistent with a randomly distributed condition and not with a simple dominant or recessive inheritance.

ACKNOWLEDGMENTS

We wish to acknowledge with gratitude the valuable help of Dr. John S. Lawrence, Director, Empire Rheumatism Council Field Unit, who participated in the examination of the respondents. We are grateful to the Division of Indian Health of the U.S. Public Health Service and the Blackfeet Tribal Council for their effective assistance and encouragement.

REFERENCES

1. Ball, J., 1950. Serum factors in rheumatoid arthritis agglutinating sensitized sheep red cells. Lancet 2: 520–523.
2. Bozicevich, J., Bunim, J. J., Freund, J. and Ward, S. B., 1958. Bentonite flocculation test for rheumatoid arthritis. Proc. Soc. Exerp. Biol. and Med. 97: 180–183.
3. Bremner, J. M., Alexander, W. R. M. and Duthie, J. J. R., 1959. Familial incidence of rheumatoid arthritis. Ann. Rheum. Dis., 18: 279–284.
4. Cotterman, C. W., 1937. Indication of unit factor inheritance in data comprising but a single generation. Ohio J. Sci. 37: 127–140.
5. Goldenberg, A., Singer, J. M. and Plotz, C. M., 1960. Hereditary factors in rheumatoid arthritis. Arthritis and Rheumatism 3: 446.
6. Lawrence, J. S. and Ball, J., 1958. Genetic studies on rheumatoid arthritis. Ann. Rheumat. Dis. 17: 160–168.

735-712 O—65——16

7. Neri Serneri, G. G. and Bartoli, V., 1956. Ricerche sui fattori ereditari del reumatismo cronico primario. Acta Genet. Med. (Roma) 5: 402–425.
8. Penrose, L. S., 1935. The detection of autosomal linkage in data which consist of pairs of brothers and sisters of unspecified parentage. Ann. Eugen. 6: 133–138.
9. Schmid, F. R. and Slatis, H., 1962. Increased incidence of serum and clinical abnormalities in spouses of patients with rheumatoid arthritis. Arthritis and Rheumatism, 5: 319.
10. Stecher, R. M., Hersh, A. H., Soloman, W. M. and Wolpaw, R., 1953. The genetics of rheumatoid arthritis: analysis of 224 families. Am. J. Human Genet. 5: 118–138.
11. Ziff, M., Schmid, F. R., Lewis, A. J. and Tanner, M., 1958. Familial occurrence of the rheumatoid factor. Arthritis and Rheumatism 1: 392–399.

PART IV—Continued

Longevity and Coronary Artery Disease

Family Patterns of Longevity and Mortality*

By Bernice H. Cohen, Ph. D., *Department of Chronic Diseases, Johns Hopkins School of Hygiene and Public Health, Baltimore, Md.*

INTRODUCTION

Among the numerous defects in scientific research, two major types seem to predominate. On the one hand, many investigations are not original—they repeat the same studies, same methods, same approaches. And, on the other hand, many researchers accept too readily, old, supposedly established concepts based on insufficient or unreliable evidence—concepts which by sheer repetition have become "axioms." On just such an insecure foundation rests our present knowledge of the role of heredity in determining life span.

Space does not permit a comprehensive review of the literature; a detailed, critical analysis will be published elsewhere. My purpose here is:

Firstly, to define the problem and the assumptions that will be considered acceptable;

Secondly, to refer briefly to some of the previous and current studies of family patterns of mortality, with particular reference to methodological problems;

Thirdly, to outline two studies currently in progress in the Department of Chronic Diseases at the Johns Hopkins School of Hygiene, and to describe not only their objectives, but also, and primarily, their procedures: how they attempt to avoid some of the pitfalls of previous studies, what problems are involved in these current inves-

*This investigation was supported in part by Public Health Service research career program award number 5-K3-GM-5590 (formerly GM-K3-5590; GSF-486) from the Division of General Medical Sciences and Public Health Service research grant number HE-06023 (formerly H-6023).

Some of the computations were done in the Computing Center of the Johns Hopkins Medical Institutions which is supported by Research Grant FR-00004 from the National Institutes of Health. In addition, since this paper was presented, support for continuation of the Special Group Study has been received from the Public Health Service (Contract No. PH-86-64-18 and later Grant No. CD-00043-01).

tigations, and what improvements in methods for future planning might be suggested; and

Lastly, to present some preliminary results from these studies.

THE PROBLEM AND ASSUMPTIONS

While it is clear that each species has its own characteristic life span and there is substantial evidence in experimental organisms for genetically determined variation within species (Pearl, 1928; Chai, 1959; Hartwig, 1959; Smith, 1959; Mühlbock, 1957; Comfort, 1956; Strong, 1936), one cannot extrapolate directly from experimental organisms to man, nor carry out studies of similar nature in man, where neither test matings nor laboratory controlled environmental conditions are possible. For the human geneticist, the direct epidemiological approach is the realistic one.

The first step in such an approach is to determine whether there are differences in human mortality that follow family patterns. If there is a genetically controlled variability within the human species (whether it involves hereditary disease susceptibilities, patterns of cellular degeneration, or somatic mutation rates), there should be familial aggregation in regard to length of life. Certainly, the demonstration of familial aggregation does not imply a genetic factor unequivocally, since this may result from nongenetic causes. Nevertheless, positive findings of familial aggregation are the usual prerequisite for evaluating the role of heredity, whereas the absence of familial aggregation would contraindicate the testing of specific genetic hypotheses, except under unusual circumstances.

PREVIOUS AND CURRENT STUDIES OF FAMILY MORTALITY

The concept of familial longevity and "brachybioty" is certainly not a new one. Many family studies of life span have been carried out, a few dating back to the 19th century (Beeton and Pearson, 1899). Some of the published studies are summarized in table 1, where they are presented in chronological order.

These reports can be classified into two broad groups: firstly, those based on *genealogy records*, including both large individual kinships (the Hyde family, the Peirce family) and collections of genealogy records (Foster's Peerage, Burke's Landed Gentry); and secondly, those based on *samples from special subgroups* of the population (e.g. workingmen's families, families of living aged persons, senescent twins, insurees, etc.).

The results of the genealogy studies, both single kinships and collections, appear to show distinct familial patterns of mortality.

There are, however, certain problems inherent in the use of old genealogy records and other defects that may have led to biases in the data. Even if the validity of recorded ages and persons is accepted (and there is no objective means of verification or reliability), the statistics have proved inadequate for other reasons, such as:

(1) the inadvertent omission of some family members, with the completeness of ascertainment varying by sex [1] and by age; [2]
(2) the pooling of family data covering several centuries without allowance for the decline in mortality rates with calendar time; *i.e.* the cohort effect (Yuan, 1931; Jalavisto, 1951);
(3) the failure to adjust for combining of different socioeconomic classes which are known to vary in mortality experience (Jalavisto, 1951); and,
(4) the absence of any reference population or comparison group for the time period studied, or any other time period.

It is difficult to estimate the effect of these biases on the reported correlation coefficients, average ages at death, proportionate mortality, and life expectancies reported from studies of genealogy records, but certainly the results would be considerably altered, and, in most cases, in the direction of indicating spurious relationships.

Studies of family mortality patterns which have involved the investigation of samples from special groups in the population, however, have not entirely eliminated all these problems either. While the use of population segments is more desirable than the investigation of genealogy records, incompleteness of data and unsystematic or irregular sampling procedures can provide biases in population studies just as serious as those found in genealogy studies. Such was the situation in Pearl's investigation of the FHR series (family history records extracted from those collected by the Department of Biology Laboraory for various other projects) and in his investigation of the families of nonagenarians, who had been ascertained through newspapers and friends (Pearl, 1931; Pearl and Pearl, 1934). Kallmann *et al.* (1956) and Kallmann's (1956) studies of senescent twins and their families, aside from questions of acertainment and zygosity diagnosis, and irrespective of the quality of the data, pertain to only a very unique segment of the population. Because twinning *per se* may be related to fertility and fertility to longevity, it is difficult to extrapolate from twin studies to the general population. There is, in fact, no

---

[1] Data recorded in Burke's Landed Gentry and Foster's Peerage (Beeton and Pearson, 1899) as well as the southern Chinese family studies by Yuan (1931) and others were almost entirely confined to males. Yet to limit the analysis to one sex eliminates perhaps critical evidence (e.g., the sex differences suggested by Jalavisto, 1951.)

[2] Many of the early records (Beeton and Pearson, 1899; Bell, 1918) showed a marked deficiency of those dying in infancy and late adult life with more complete reporting of young adult deaths. As a consequence, distributions of deaths show marked distortion when compared with vital statistics of the time (Carlisle's 18th century life tables and Glover's life tables for Massachusetts males 1890). Inclusion of all deaths irrespective of year of birth relative to the date of conclusion of study period contributed at least in part to this artificial distribution.

TABLE 1.—*Family studies of mortality and life span*

Investigators	Date published	Source of data	Time period covered	Classification of data	Method of analysis and results
Beeton and Pearson	1899 1901	a) Old English Genealogy Records: 1) Foster's Peerage of Great Britain & Ireland. 2) Burke's Landed Gentry of the British Empire. 3) Society of Friends Records. b) Family Data—Colleagues at Univ. College.	Dating from 17th century to end of (?) 19th century.	Genealogy records—upper social classes, confined mainly to males. Same as above. Genealogy records of small traders & yeomen. Histories reported by family members—professional classes mainly.	Correlation coefficients. Results: Positive correlation between length of life: parents—children and pairs of sibs. Greater similarity between collaterals, but in general, low coefficients with some negative.
Bell	1918	The Hyde Family Descendants of William Hyde who died 1681 in Norwich, Conn. Publ. 1864.	Mid-1600's to 1864	A single genealogy	Percentages, distributions, average duration of life for offspring of various parental combinations and for parents of offspring of different classification. Results: An almost direct relationship between life span of offspring and parents noted.
Wilson and Doering	1926	Peirce family descendants in male line from John Pers. Publ. 1881.	1637–1880	Single family genealogy	Correlation coefficients; distributions; and average ages of death for offspring of different parental combinations. Analysis confined to those who lived to be 20 years and could have reached 95 by end of study. Results: Father-son combinations (confined to those sons who were also fathers) yielding high correlation. Most of correlation coefficients higher than in other studies.
Holmes	1928	Allström's Dictionary of Royal Lineage.	"Earliest times to 20th century."	Genealogy records of royalty	Correlation coefficient methods similar to Beeton and Pearson. Separation by centuries. Results: Low correlation coefficients.
Pearl	1924, 1931 1931	Five American Genealogical Works; Whitney Family, Hyde Family, Spooner Family, Smedley Family, Durham Town History. Family History Records (Working men's families in Balto. & surrounding area) referred to as FHR series.	1678–1913 1800's to 1920's?	Genealogy records Personal interview data concerning families of working men.	Correlation coefficients. Results: No significant deviation from "zero" correlation. Life tables—standard actuarial methods. Results: Greater life expectancy for parents of children dying over 50 than for those of parents dying under 50. Mean after life times of children were differentiated from one another in accordance with age of death of parent.

	1934	Family Histories of Nonagenarians.	Probably late 1700's to 1920.	Data on families of nonagenarians (mailed questionnaire).	Life tables. TIAL=total immediate ancestral longevity (sum of ages at death of 2 parents, 4 grandparents) higher in nonagenarians than for younger living propositi of FHR series.
Yuan	1931	Chinese Genealogy	1365–1914	Single family genealogy	Life table functions. Survivorship curves (l_x); expectation of life (e_x^0) and probability of dying (q_x). All persons born after 1819 and dying before 20 years of age omitted. Results: Small amount of material, small differences, but consistency in that parental life expectancy greater among long lived sons and sons' life expectancy greater if parents were long lived.
Stoessiger	1932	Pearson's Family Data and families of Macpherson's private patients.	19th century primarily?	Collected families: 2/3: Pearson's family data (i.e. genealogy records). 1/3: family history interview data from Dr. Macpherson's private patients, friends and colleagues in Glasgow.	Reanalysis of Pearson's and Macpherson's data excluding deaths from accidents or infection—correlation coefficient method. Results: Similar to those of Beeton and Pearson.
Kallmann	1948 and ff	Twin pairs—old aged homes, institutions, population.	19th century to present.	Senescent twin pairs and their families.	Twin analysis—mean intrapair differences in survivorship of MZ vs DZ twins. Mean life spans of sibs of twins classified by parental age at death. Results: Mean intrapair life span differences greater for DZ than for MZ. Sib life span increased with those of parents.
Jalavisto	1951	Genealogy Records Finnish and Swedish Middle Class & Nobility.	1500–1829	Old genealogy records of numerous families.	Mean ages of death of offspring of parents with different life spans. Results: Both maternal and paternal longevity associated with mean life span of offspring. Maternal effect greater than paternal with daughters showing practically no paternal effect.
Insurance Studies: Specialized Mortality Investigation.	1903	Actuarial Society of America (based on 34 life companies) involving 300,000 insured men. 1869–1899.	? to 1899	Insured persons and their families.	Death rates at successive ages among male policy holders classified according to longevity of parents. Mortality of men whose parents lived to 75 was lower than those whose parents died under 60. Life expectancy greater for those with long lived parents.

TABLE 1.—*Family studies of mortality and life span*—Continued

Investigators	Date published	Source of data	Time period covered	Classification of data	Method of analysis and results
Symonds	1913	Mutual Life Ins. Co., 116,590 men insured 1891–1899 traced to 1909.	? to 1909	Insured persons and their families.	Mortality experience using a percent scale based on average family history (including grandparents). Results: Mortality of those with most favorable family history 25 percent better than those with unfavorable one.
Hunter	1932	New York Life, 314,000 persons associated with standard male risks insured 1914–1917 traced to 1931.	? Mid-1800's ? to 1931.	Insured persons and their families.	Mortality experience (ratio of actual to expected deaths) of policy holders grouped by rating scheme for good, average, poor, family histories. Results: Small variation in 3 groups but mortality ratios were slightly better among those with good family histories at all ages up to 50 years at policy issue.
Marshall	1932 (1929 study pub. 1932).	Provident Mutual 30,000 policies issued Dec. 1908–Aug. 1912 traced to 1928.	? Mid-1800's ? to 1928.	Insured persons and their families.	Mortality experience (ratio of actual to expected deaths) of insured men and their sibs classified by parental history. Results: Mortality rate lower among insurees both of whose parents were living at policy issue than when parents dead.
Dublin & Marks	1941–1942	Metropolitan Life Ins. Co., 118,000 insured white men, Ordinary Dept. living 20–64 in 1899–1905 traced to 1939.	? to 1939	Insured persons and their families.	Ratio of actual to expected deaths of insured men grouped by parental living-dead status and age at death. Results: Survivorship of insurees associated with living status of parents at time of issue and apparently independent of age at death when both parents dead.

meaningful reference population in any of these studies of special subgroups.

In contrast to all of the investigations which have indicated positive familial associations in regard to life span, some of the insurance company studies suggest a relationship of the mortality experience of insurees with the living or dead status of their parents at the time of policy issue, irrespective of parental age at death. As indicated by Dublin and Marks (1941), however, there are many defects in life insurance material, including "an appreciable amount of misstatement, intentional or unintentional, of ages and causes of death" of family members, with no attempt to verify these statements on the part of insurance companies. There is also the definite bias introduced by medical selection of persons for insurance—a selection often influenced by family history, and thus leading to more rigid screening on physical criteria for those with poor family mortality experience.

Moreover, where no actual body of recorded data was utilized either as a source, or for verification [as was probably the case in Pearl's FHR and nonagenarian series as well as Pearson's family data from colleagues, Stoessiger's collected families (1932–33), the insurance data and others], human longevity reports have depended largely on "unsupported memory and tradition in a field where the emotional premiums of exaggeration are high" (Comfort, 1956).

Thus, though there are much data on the familial patterns of life span, their interpretation is questionable, since the suggested positive relationships as well as any lack of relationships are all explicable by one bias or another that is involved in the selection, completeness, or the source of the data.

PLAN OF STUDIES IN PROGRESS

In view of this situation, two studies of patterns of mortality and aging were initiated in the Department of Chronic Diseases at the Johns Hopkins School of Hygiene to determine: firstly, whether there are family patterns of mortality; secondly, if they do exist, the nature of those patterns in terms of age at death, cause of death, and cause-age interaction; and thirdly, to identify in living subjects possible indicators related to family mortality patterns.

The Population Study:

For the first and main study, a sample of deaths (deceased probands) occurring in Baltimore City to U.S.-born white residents has been selected and is being matched with living controls of the same sex, race and neighborhood, and U.S. born within 5 years prior to the deceased probands. Detailed personal and family data are being obtained concerning each proband and each living control. The mortality experience of relatives (mothers, fathers, sisters, brothers,

spouses and spouses' families) of the deceased probands will be compared with that of the corresponding family members of the matched controls. Intrafamilial as well as interfamilial mortality patterns will be examined. The influence and interaction of various biological and social factors (such as size of sibship, birth order, parental age at birth of subject and number of offspring, as well as number of marriages, divorces, occupation, education, religion, etc.) will be evaluated for a better comprehension of the relationship of the extrinsic to the intrinsic factors. To gain an insight into the possible effect of assortative mating on familial mortality experience, the mortality of spouses and family members of spouses of dead probands will be compared with spouses and corresponding family members of living controls.

To obtain patterns characteristic of the general population, a sample of *total deaths of the community* was selected. It was estimated that a sample of deaths and a sample of living controls in excess of 520 each would be required for the detection of certain reasonable differences. [Samples of 520 should detect at least a 10 percent difference between dead probands and living controls in the percentage of dead mothers, dead fathers, sisters and/or brothers at the 5 percent level of significance with a 90 percent probability, even if the smaller percentage is 40 percent (*i.e.*, assuming a level of maximum standard error). When total sibs are pooled (on the assumption that the number of sibs would be similar to that observed in a pilot study described below) a difference of 5 percent should be detected at a 5 percent level of significance with 80 percent probability. Pooling of parents still assures detection of only a 10 percent difference but with 95 percent probability at the 1.0 percent level of significance.]

Thus with allowance for an attrition rate of 5 percent due to refusals, inability to locate, and other factors, a sample of 550 deceased probands and 550 living controls would be adequate (to detect differences at the above levels). These are only approximations, however, since, in the analysis, life table factors will be used, while the estimates of standard errors of the above are based on binomial percentages.

The sample of deaths stratified by age was selected from the 6,535 deaths in Baltimore City in 1960 meeting the specifications for eligibility: U.S.-born white residents of Baltimore City and/or Baltimore County (table 2). In each of the age groups indicated, 110 deaths were selected by a combined random number and systematic sampling procedure with the method of selection taking into account seasonal variation in deaths. Table 3 shows the distribution of the sample of deaths according to age and seasonal quarter of year. Table 4 presents the death certificate causes of death of the proband sample as compared with distribution of causes of death for Baltimore City residents in 1960.

TABLE 2.—*Distribution of Baltimore deaths and selected sample by age group (1960)*

Age group	Deaths of Baltimore City and Baltimore County residents in Baltimore City		Eligible* 1960 deaths	Sample size
	Nonwhite	White		
<20	605	687		
20–39	336	356	297	110
40–49	468	622	536	110
50–59	646	1,349	1,127	110
60–79	1,280	4,925	3,467	110
80+	211	1,746	1,108	110
Total	3,546	9,685	6,535	550

*Eligible=meeting specifications of U.S.-born white residents of Baltimore City and/or County, over 20 years of age.

To minimize spurious correlations within families due to heterogeneity of socioeconomic groups, the living control group consists of the nearest neighbors of the deceased probands matched on the criteria already indicated. A systematic method for selecting such neighborhood controls was developed, including maps, forms, and a detailed outline of procedure. The neighborhood control selection sheet as shown in figure 1 gives the necessary data for matching, and space for the screening record. The Neighborhood Control Selection Route (figure 2) shows the walking pattern to be followed in selecting a city neighborhood control. The proband's household is marked with an "X" in Block 1.

A sample of a control selection route mapped for a "city" control is given in figure 3. Note the indication of census tract boundaries. In Baltimore City, all controls are being matched within the census tract of the proband. If a census tract boundary intervenes, it becomes a stopping point for screening in that direction. Then the searching pattern is enlarged, fanning out in other directions within the proband's census tract.

For county screening, election district boundaries are used, instead of census tract boundaries. An outline map of the whole county divided into election districts (figure 4), as well as a map of the election district shape and boundaries for the specific proband to be matched, are provided with each control selection sheet for county controls. A sample "county" control is mapped in figure 5—the "X" indicates the location of the proband household.

Variation in socioeconomic factors not taken care of through matching by place of residence is to be evaluated by the analysis.

Another pitfall in some studies, the influence of secular trends observed in investigations encompassing large spans of calendar time, should be made negligible by means of the matching of probands with controls of similar cohorts.

To avoid incompleteness and selection bias, *all* family members are being included, whether living or dead. Also, several respondents are

TABLE 3.—*Sample selection—1960 deaths to U.S.-born white residents of Baltimore City and/or County dying in Baltimore City, selected by quarters of year*

(Jan.-Mar., Apr.-June, July-Sept., Oct.-Dec.)

Based on total observed deaths to eligibles—1960

Age group	Total eligible deaths in 1960	January-March				April-June				July-September				October-December			
		Dx fraction × 110			Second adjusted sample 3/27/61	Dx fraction × 110			Second adjusted sample 3/27/61	Dx fraction × 110			Second adjusted sample 3/27/61	Dx fraction × 110			Second adjusted sample 3/27/61
20-39	297	85	.2862	31.48	31	72	.2424	26.66	27	64	.2155	23.71	24	76	.2559	28.15	28
40-49	536	157	.2929	32.22	32	112	.2090	22.99	23	123	.2295	25.25	25	144	.2687	29.56	30
50-59	1,127	311	.2760	30.36	30	290	.2573	28.30	28	247	.2192	24.11	24	279	.2476	27.24	27(28*)
60-79	3,467	1,019	.2939	32.33	32	811	.2339	25.73	26	751	.2166	23.83	24	886	.2556	28.12	28
80+	1,108	344	.3105	34.16	34	229	.2067	22.74	23	229	.2067	22.74	23	306	.2762	30.38	30
Total	6,535	1,916			159	1,514			127	1,414			120	1,691			143(144)

*One extra needed for total 110, added.

TABLE 4.—*Comparison of proband causes of death with reported causes of death in Baltimore City (1960), Baltimore City Health Department Annual Report*

Age	Source	Infectious diseases		Neoplasms		CVRD		Accidents		Suicides		Diabetes		Others		Total	
		No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent
20–39	D.P.*	8	7.27	23	20.90	37	33.63	17	15.45	8	7.27	1	0.91	16	14.54	110	20.00
	Balto.**	20	7.63	53	20.22	74	28.24	44	16.79	21	8.01	6	2.29	44	16.79	262	3.44
40–49	D.P.*	4	3.64	27	24.54	48	43.63	8	7.27	2	1.82	1	0.91	20	18.18	110	20.00
	Balto.**	38	7.81	121	24.89	200	41.15	33	6.79	14	2.88	3	0.61	77	15.84	486	6.38
50–59	D.P.*	5	4.54	33	30.00	54	49.09	2	1.82	0	0.00	3	2.73	13	11.81	110	20.00
	Balto.**	70	6.50	262	24.34	547	50.83	38	3.53	13	1.21	17	1.58	129	11.98	1,076	14.14
60–79	D.P.*	2	1.82	20	18.18	67	60.90	0	0.00	1	0.91	5	4.54	15	13.63	110	20.00
	Balto.**	202	4.79	789	18.71	2,639	62.59	77	1.83	27	0.64	133	3.15	349	8.27	4,216	55.42
80+	D.P.*	5	4.54	11	10.00	82	74.54	4	3.63	0	0.00	2	1.82	6	5.45	110	20.00
	Balto.**	68	4.34	139	8.87	1,217	77.66	41	2.61	2	0.13	22	1.40	78	4.98	1,567	20.59
Total	D.P.*	24	4.36	114	20.72	288	52.36	31	5.63	11	2.00	12	2.18	70	12.72	550	100.00
	Balto.**	398	5.23	1,364	17.93	4,677	61.48	233	3.06	77	0.10	181	2.38	677	8.90	7,607	100.00

*Deceased Probands. **Reported causes of death for Baltimore City, including foreign born.

NEIGHBORHOOD CONTROL SELECTION SHEET

Proband

Reg. No. _____ Sex: Male _____ Female _____

Address: _____ Birth Date: _____ Month/ Year

Death Date: _____ Month/Day/Year

Control

Sex: Male _____ Female _____ Birth Date: From _____ to _____ Month/Year Month/Year

Age: From _____ To _____ Census Tract: _____

Street Number	Telephone Number	Full Name	Comments	Ages Male	Female

FIGURE 1.—Neighborhood control selection sheet.

usually interviewed for each family to prevent omissions of family members—even those dead in infancy.

Besides provisions for assuring completeness of records, *validation procedures* for assuring accuracy are incorporated. To verify the ages of the living, households of all living relatives are being contacted, and to verify ages at death and causes of death for deceased relatives, death certificates and other vital records are obtained.

The types of data being collected are outlined in *appendix A*.

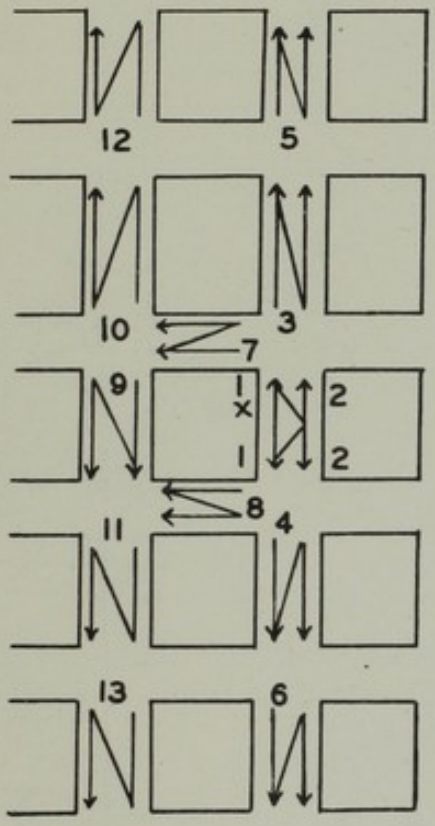
In addition to the detailed information obtained on the biological relatives of the probands and controls for comparison of their respective mortality patterns, the same types of data are recorded on spouses and spouses' relatives in order to examine the role of assortative mating and adult environment shared by the spouses as compared to the influence of genetic factors and childhood environment shared by the sibs.

MAP FOR CONTROL SELECTION ROUTE

Reg. No. _____

Each square equals one square block

X equals home of proband



Check blocks 1-13 to determine whether they are within same census tract. If so:

1. Proceed in numbered order

2. When starting a new block (3), begin with block closest to proband's home.

3. If a dead end or wooded area interrupts the normal order, proceed to the next consecutive number

4. If there is no answer at a home, mark this on Neighborhood Control Selection Sheet for a return call. Proceed onward and obtain control. Mark this control "Tentative Control" until all closer dwellings are checked out.

   If on a return call, a control is found closer to the home of the proband, he will be the control, and the "Tentative Control" will be removed or used as a secondary control, according to a subsequent standardized routine for secondary controls.

5. If no control has been found at the end of block 13, return for further instructions.

If any of the blocks fall outside the census tract of the deceased proband, cross these out and outline the area of the pattern within the given census tract.

Proceed with only those blocks within the census tract in their numbered order.

FIGURE 2.—Map for neighborhood control selection route.

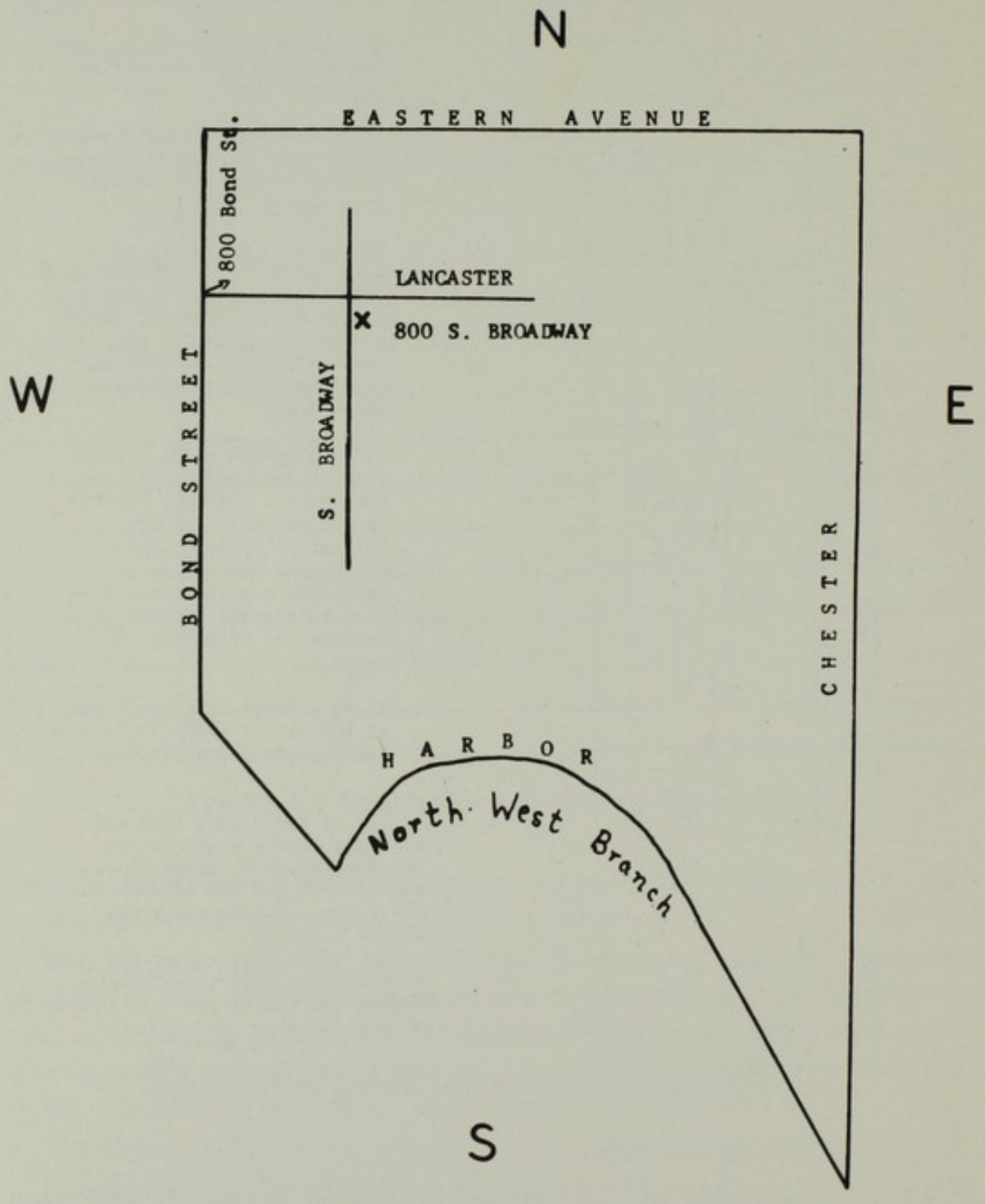
Progress of Population Study:

Actual scheduling of the various aspects of the study has been subject to revision according to the relative progress of different phases. A brief step-by-step summary of the investigation since its inception may help clarify the application of some of the procedures under discussion.

Before any work was begun, a *pilot study* was carried out on respondents for a sample of 50 consecutively recorded Baltimore City

735-712 O—65——17

CITY CONTROL P 11264



PROBAND LIVED AT 800 S. BROADWAY

FIGURE 3.—Sample map for a "city" control. "X" marks location of proband's residence.

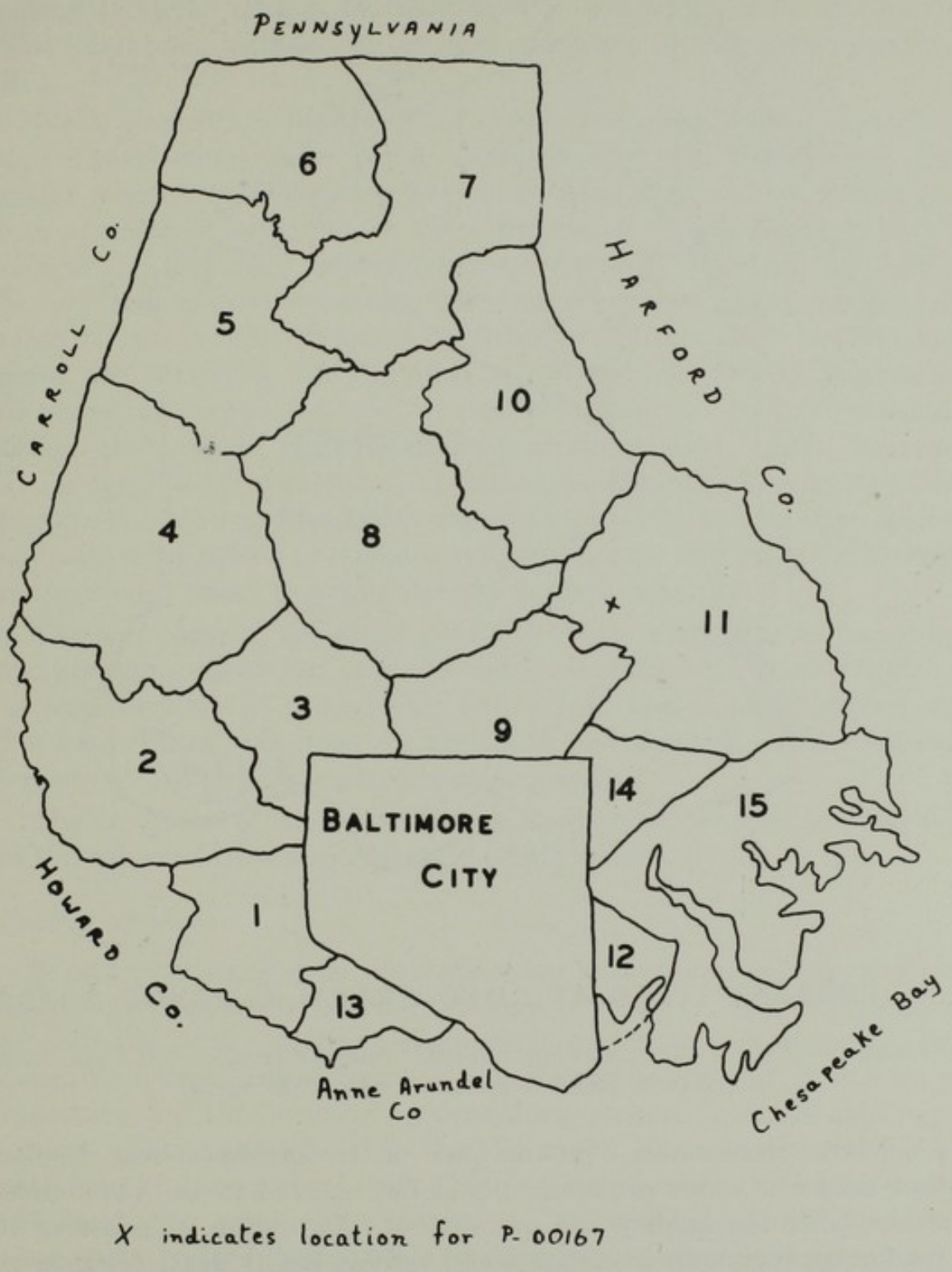


FIGURE 4.—Map of Baltimore county election districts as they surround Baltimore.

deaths occurring in March 1959 to white persons over 20 years of age. A trial interview schedule was developed for this purpose. The problems involved in tracing, interviewing, questions, etc., were noted, and suitable adjustments in the interview schedules and

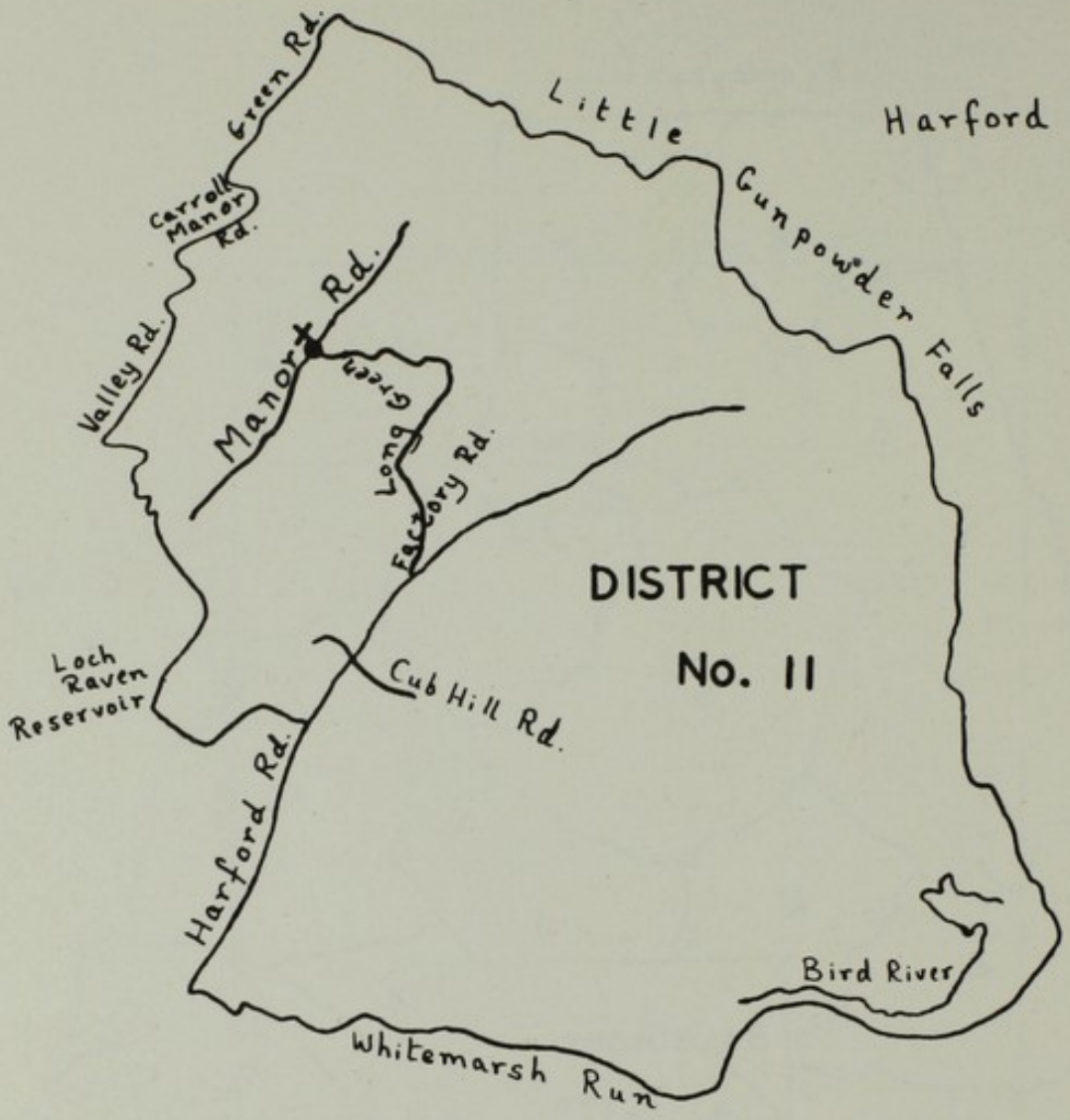


FIGURE 5.—Map for a sample county control. Proband's residence at Long Green and Manor Roads is indicated by an "X" mark.

procedures were made. On the basis of the findings it was decided that only U.S.-born probands would be included in the population study, since the quality and quantity of information on relatives of the foreign born were inadequate and verification by death certificates not feasible, whereas the data obtained from relatives of native-born whites were quite satisfactory.

In this pilot study, information was collected on 593 persons derived from families of 38 dead persons (12 probands having been rejected for various reasons, such as place of residence, refusal, etc.),

thus providing an estimated 15.6 persons per proband to be expected in the family study—the basis for determination of the required sample size as already indicated.

After the pilot study, the main investigation was initiated in the spring of 1961. It can be considered as consisting of three phases. The operations included in the *first phase* are listed in the appendix B–1.

While procedures for the total study (including control matching and interviewing) were being established in this first phase, the major areas of concentration were in revising the interview schedules into their final form and in the selection of the deceased proband sample, a necessary prerequisite to control matching. Priority in time was advisable for the collection of proband data for several reasons. Since the deceased proband's address was the starting point for searching for respondents, it was considered important to make the contact as soon as possible after the proband's death before the nearest relatives moved away or died. Also, because living controls were matched to deceased probands, it was desirable to have deceased proband data concerning such factors as exact home address, transient or permanent resident status, and actual inclusion in the study, before searching for the control match.

The *next phase* of the study continued with many aspects begun in the preceding phase but here the emphasis shifted (1) from initiating to final revision of procedures; and (2) to the primary objective: the major data collection on probands and controls. The various aspects of this phase are listed in appendix B–2.

The *last phase* of the study (appendix B–3) will consist of analysis of the collected data and reporting of the results.

Special problems:

Several methodological problems were encountered which may be of interest to those considering this type of study. Though matching of probands with "nearest-neighbor" living controls is theoretically desirable to minimize differences in socioeconomic status, it is a very expensive and time-consuming procedure, especially in the older age groups. Our statistician, Dr. Earl Diamond, has recently calculated that for a 50-percent probability of obtaining a match for an 80-year-old male, one would need to screen 90 households in Baltimore. Consequently, for future studies where suitable population controls are needed, Dr. Diamond has suggested a "balancing" procedure, the details of which will be published. In the revised design, a random number type of household sampling is used where the sampling probabilities are based on the age, sex and neighborhood distribution of the proband population to be matched. This method, therefore, results in the selection of a control group which has the same age,

sex and neighborhood composition as the proband group without individual proband-control matching.

Another difficulty encountered has been locating respondents for elderly deceased probands and determining the exact dates and places of death of their relatives, especially when the index case was the last surviving member of the family or one of the few surviving members, the others being ill or senile. Moreover, even when suitable family respondents are available, often some of the relatives died so long ago that the recall of survivors for exact details is poor and/or the official records are inadequate.

The handling of death certificate verification locally and out of state for the proband and control families has developed into a sizeable aspect of the study. For obtaining death certificates, a minimum amount of accurate information is required. Often contacting the cemetery of burial to obtain the exact date of death and other information has proved helpful. The degree of searching that each state will carry out around approximate dates depends upon its staff, facilities and the organization of its vital statistics divisions. These are highly variable from state to state. Moreover, even within individual states—and Baltimore City itself—the filing system for death records has changed numerous times in the past century.

When death certificates are not available from state health departments, they can sometimes be obtained from local county, city or borough divisions. Some useful pamphlets include the Rand McNally Geographical Handbook, a keyed index of the United States locating counties and all cities and towns of over 100 population, the Cemetery Directory of the American Cemetery Association and "Where to Write for Birth and Death Records, United States and Outlying Areas" PHS Publication No. 630 A–1, to mention a few.

Another problem not to be overlooked in studies where respondents have entered the sample without their own knowledge, interest or motivation is that of obtaining cooperation and maintaining rapport.

Special Group Study:

The second investigation of family mortality is very different in design from the first. Planned to gain clues concerning manifestations of family mortality patterns in living subjects, it deals with test responses, measurements and family mortality experience in a special group. This special group comprises about 500 males from 24 to 100 years of age who have volunteered to participate in a longitudinal study of aging being carried out by the Gerontology Branch of the National Heart Institute. About every 18 months these subjects spend 2½ days at Baltimore City Hospitals going through a series of physiological, biochemical, and psychological tests and anthropometric measurements. In addition, our staff is obtaining the following:

(1) Personal and family history data: Information similar to that being collected in the population study just described is being obtained by means of a combined interview and self-administered questionnaire.
(2) Smoking histories comparable to those used by Haenszel, Shimkin and Miller (1956) are being filled out by subjects with the assistance of interviewers.
(3) Ability to taste phenylthiourea (PTC) is being tested.

The results of the test responses, anthropometric data and the PTC tests, smoking and other personal data will be correlated with each other and with the sociobiological factors as well as with the family mortality data in order to gain clues as to possible indicators related to family mortality.

Although a volunteer group is not the most satisfactory type of sample, nevertheless, the detailed longitudinal study of this cooperative and well-motivated group makes available a unique combination of data which would be unobtainable in the larger, community-based investigation. A less than ideal approach, therefore, seems justifiable to gain possible insight into the "natural history of mortality patterns," an area where there is very little concrete information at this time.

PRELIMINARY FINDINGS

Both of the studies discussed here, the population study and the special group study, are in the data collection phase. However, some preliminary findings are available for the population study, though they are limited to an analysis of the first 101 matched pairs (101 probands and 101 controls). Moreover, these results must be considered with caution since they constitute less than 20 percent of the sample, with the information not yet totally verified.

In religious affiliation, probands and matched controls show almost identical distributions (table 5). On the other hand, levels of education differed slightly. As shown in table 6, more probands had been graduated from high school or gone beyond high school than controls, with the difference residing in those dead under 50 years of age, (i.e. born after 1910) and no difference in those over 50 years of age.

TABLE 5.—*Religion of probands and controls*

	Probands		Controls	
	No.	Percent	No.	Percent
Protestant	47	46.5	49	48.5
Roman Catholic	36	35.6	37	36.6
Jewish	16	15.8	15	14.9
Other	1	.99	----	----
None	1	.99	----	----
Total	101	----	101	----

TABLE 6.—*Education of probands and controls*

Age	<50				50+				Total			
	Probands		Controls		Probands		Controls		Probands		Controls	
	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent	No.	Percent
High School and beyond	40	56.3	31	43.7	7	23.3	7	23.3	47	46.5	38	37.6
Less than High School Graduates	30	42.3	40	56.3	23	76.7	23	76.7	53	52.5	63	62.4
Unknown	1	------	------	------	------	------	------	------	------	------	------	------
Total	71	-----	71	------	30	------	30	------	101	------	101	------

Regarding parental survivorship, as shown in table 7 and figure 6, there were no differences between parents of probands and parents of controls in survival to 50 years of age. Fathers of probands had the same life table probability of surviving to 50 as fathers of controls (.739 and .740 respectively); and similarly the mothers of probands and mothers of controls (.837 and .818 respectively). There is, however, a suggestion of a lower probability of surviving to 70 years of age for fathers of probands (.323) than for fathers of controls (.393) with again a similar trend—but smaller difference—for mothers (.527 versus .571).

Probands and controls had the same proportion of total living fathers (26/101, 25.7 percent, and 25/101, 24.8 percent, respectively) with probands having a slightly higher proportion of fathers living under 70 than controls (19.8 percent versus 15.8 percent). Of

TABLE 7.—*Survivorship of parents and sibs of probands and controls*

	Fathers		Mothers	
	Probands	Controls	Probands	Controls
Total Number	101	101	101	101
Cum P_{50}	.739	.740	.837	.818
S.E.	±.062	±.062	±.052	±.055
Number reaching 50 years	85	86	88	88
Number alive under 50	3	2	5	4
Cum P_{70}	.323	.393	.527	.571
S.E.	±.050	±.053	±.057	±.057
Number reaching 70 years	25	35	35	47
Number alive under 70	20	16	30	21
	Brothers		**Sisters**	
	Probands	Controls	Probands	Controls
Total Number	190	194	187	188
Cum P_{50}	.700	.827	.806	.827
S.E.	±.037	±.031	±.033	±.031
Number reaching 50 years	65	80	64	76
Number alive under 50	80	88	96	87
Cum P_{70}	.505	.579	.722	.697
S.E.	±.057	±.056	±.050	±.052
Number reaching 70 years	3	11	9	6
Number alive under 70	131	141	147	150

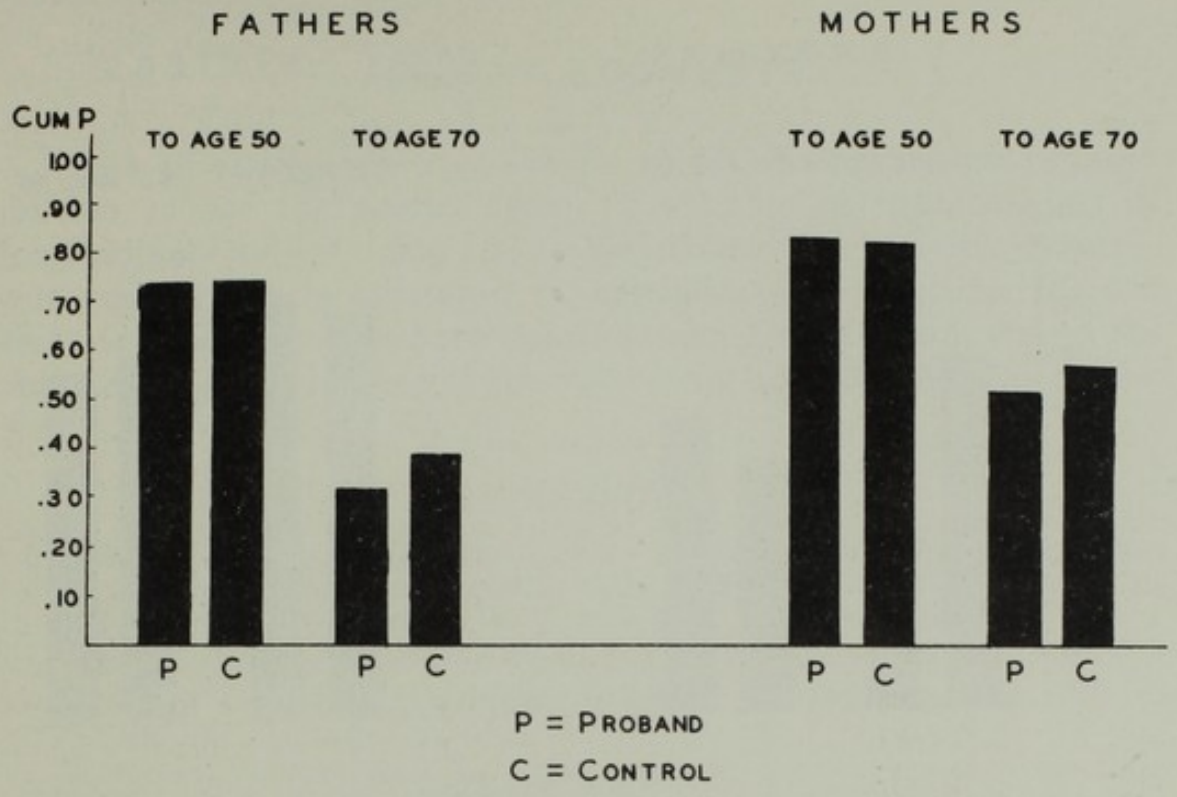


FIGURE 6.—Life table probabilities of surviving to age 50 years (Cum P_{50}) and to age 70 years (Cum P_{70}) for fathers and mothers of probands and controls.

course, very few living fathers were under 50. On the whole, probands had slightly more living mothers (48/101, 47.5 percent, versus 40/101, 39.6 percent), and thus more living mothers under 70 years (30/101, 29.7 percent) than controls (21/101, 20.8 percent).

As for sibs, no real differences were noted in numbers of sibs of probands and controls. Probands had 190 male sibs and 187 female sibs as compared to 194 male and 188 female sibs for controls (table 7). Of these, slightly fewer male sibs of probands were alive and under 70 years of age than male sibs of controls (68.9 percent and 72.7 percent respectively). As for female sibs, probands had more living sisters under 50 than controls (51.3 percent versus 46.3 percent) though probands and controls had the same percentage alive under 70 years (78.6 percent versus 79.8 percent).

Sib survivorship experience followed a pattern similar to that of parents. As shown in table 7 and figure 7, the male sibs of probands had a significantly lower probability of surviving to 50 years than those of controls ($.700 \pm .037$ as compared to $.827 \pm .031$) and possibly less likelihood of surviving to 70 years (.505 versus .579) although the latter group involved very small numbers.

Among female sibs, there were no real differences; probability of survival to 50 years being .806 and .827 for female sibs of probands

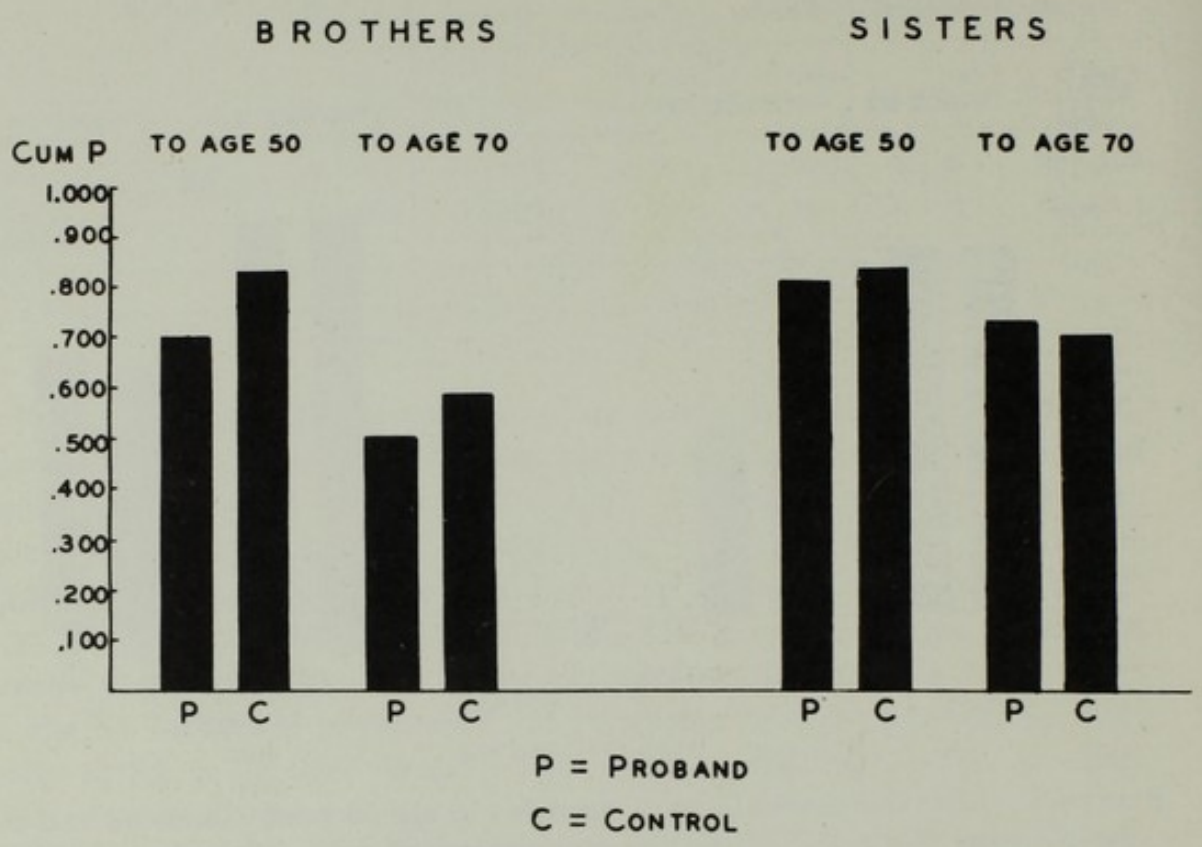


FIGURE 7.—Life table probabilities of surviving to age 50 years (Cum P_{50}) and to age 70 years (Cum P_{70}) for brothers and sisters of probands and controls.

and controls, respectively. The numerical reversal in direction for survivorship to 70 years (.722 versus .697), though appearing to be inconsistent with the other findings, involved only 9 and 6 persons, respectively.

It appears that whatever suggestion there is of survivorship differences between corresponding family members of probands and controls, these are limited to the male relatives. The female relatives of probands seem to be similar in survivorship experience to those of the controls.

These findings, however, provide only a preliminary glimpse of a very small part of the data. It is not possible to predict whether the same trends will appear when all the validated information is included and more refined statistical techniques are utilized. At this stage of the study it would not be safe to venture any conclusions nor speculations. It is hoped, however, that both the population and the special group studies may ultimately yield clues as to hypotheses (genetic or other) for further, more definitive types of investigation.

Since such studies require sizable samples, and preferably a variety of population segments (geographical, ethnic, urban, rural, etc.), more and larger population samples, perhaps collaborative projects at

several research centers, are required both for the retrospective family data analyses and for the longitudinal followup of living subjects under study.

ACKNOWLEDGMENTS

I wish to express my appreciation to Dr. Nathan Shock for permission to use the special group of subjects participating in his longitudinal study of aging and to Mr. Arthur Norris for his cooperation and invaluable assistance in administrative problems. I also want to thank Dr. Abraham Lilienfeld and Dr. Margaret Bright for their most careful reading of the manuscript and their helpful suggestions.

REFERENCES

1. Actuarial Society of America, 1903. *Specialized Mortality Investigation Experience of 34 Life Companies upon 98 Special Classes of Risks.* New York. (Cited by Dublin, L. I., Lotka, A. J., and Spiegelman, M. 1949. *Length of Life*, rev. ed. New York: Ronald Press.)
2. American Cemetery Association Bulletin. *Cemetery Directory.* 329 E. Broad St., Columbus 15, Ohio.
3. Beeton, M., and Pearson, K., 1899. Data for the problem of evolution in man. II. A first study of the inheritance of longevity and selective death rate in man. *Proc. Roy. Soc.* 65: 290–305.
4. Beeton, M., and Pearson, K., 1901. On the inheritance of the duration of life, and the intensity of natural selection in man. *Biometrika* 1: 50–89.
5. Bell, A. G., 1918. *The duration of life and conditions associated with longevity. A study of the Hyde genealogy.* Washington: Judd & Detweiler, Inc. 57 pp. (Study based on R. H. Walworth, "Genealogy of the Hyde Family" 1864.)
6. Chai, C. K., 1959. Life span in inbred and hybrid mice. *J. Hered.* 50: 203–208.
7. Comfort, A., 1956. *The biology of senescence.* New York: Rinehart & Co.
8. Diamond, E., 1963. Personal communication.
9. Dublin, L. I., and Marks, H. H., 1941. The inheritance of longevity—a study based upon life insurance records. *Tr. A. Life Insur. M. Dir. America* 28: 89–120.
10. Haenszel, W., Shimkin, M. B., and Miller, H. P., 1956. *Tobacco smoking patterns in the United States.* Public Health Monograph No. 45. PHS Publication No. 463. Washington, D.C.: Government Printing Office. (Issued concurrently with Nov. 1956 issue *Publ. Health Rep.* 71, No. 11, May 1956.)
11. Hartwig, W., 1959. Lifespan of cattle and horses under various climatic conditions and the reasons for premature culling. In: *The Lifespan of Animals* (Ciba Foundation Colloq. on Ageing, Vol. 5), pp. 57–71.
12. Holmes, S. J., 1928. Age at parenthood, order of birth, and parental longevity in relation to longevity of offspring. *Univ. California Pub. in Zoology* 31 (15): 359–375.
13. Hunter, A., 1932. Note on the effect of family history and longevity. *Tr. Actuarial Soc. Amer.* 33 (87–88): 405–408.
14. Jalavisto, E., 1951. Inheritance of longevity according to Finnish and Swedish genealogies. *Ann. Med. Int. Fenn.* 40: 263–274.

15. Kallmann, F. J., 1956. The genetics of aging from the neurologic and psychiatric aspects of the disorders of aging. *Proc. Assoc. Res. Nerv. & Ment. Dis.* 35: 95–111.
16. Kallmann, F. J., Aschner, B. M., and Falek, A., 1956. Comparative data on longevity, adjustment to aging and causes of death in a senescent twin population. *Novant' Anni delle Leggi Mendeliane*, pp. 330–339.
17. Marshall, E. W., 1932. Parental history and longevity. *Tr. Actuarial Soc. Amer.* 33 (87–88): 395–404.
18. Maurizio, A., 1959. Factors influencing the lifespan of bees. In: *The Lifespan of Animals* (Ciba Foundation Colloq. on Ageing, Vol. 5), pp. 231–246.
19. Mühlbock, O., 1957. The use of inbred strains of animals in experimental gerontology. In: *Methodology of the Study of Ageing* (Ciba Foundation Colloq. on Ageing, Vol. 3), pp. 115–130.
20. Pearl, R., 1924. Preliminary account of an investigation of factors influencing longevity. *J.A.M.A.* 82: 259–264.
21. Pearl, R., 1928. *The rate of living.* New York: Alfred A. Knopf.
22. Pearl, R., .1931. Studies on human longevity. IV. The inheritance of longevity. Preliminary report. *Human Biology* 3: 245–269.
23. Pearl, R., and Pearl, R. D., 1934. *The ancestry of the long-lived.* Baltimore: Johns Hopkins Press.
24. *Rand McNally Geographical Handbook.* 1962. A keyed index of the United States locating counties and all cities and towns of over 100 population. New York: Rand McNally & Co.
25. Smith, J. M., 1959. The rate of aging in *Drosphila subobscura.* In: *The Lifespan of Animals* (Ciba Foundation Colloq. on Ageing, Vol. 5), pp. 269–281.
26. Stoessiger, B., 1932–1933. On the inheritance of duration of life and cause of death. *Ann. Eugen.* 5: 105–178.
27. Strong, L. C., 1936. Production of the CBA strain of inbred mice: long life associated with low tumor incidence. *Brit. J. Exptl. Pathol.* 17: 60–63.
28. Symonds, B., 1913–15. Some studies in family history. *Proc. Life Ins. M. Dir. America* 23: 4–57
29. Wilson, E. B., and Doering, C. R., 1926. "The Elder Peirces." *Proc. Nat. Acad. Sci.* 12: 424–432.
30. Yuan, I-Chin., 1931. Life tables for a southern Chinese family from 1365 to 1849. *Human Biology* 3: 157–179.
31. Yuan, I-Chin., 1932. A critique of certain earlier work on the inheritance of duration of life in man. *Quart. Rev. Biol.* 7 (1): 77–83.
32. *Where to Write for Birth and Death Records, U.S. and Outlying Areas.* PHS Publication. No. 630A1 and others.

APPENDIX

Community Based Population Study

A. *Types of data being collected*

1. *On dead probands only*
   Complete death certificate information including age at death, cause of death, year of death, place of burial, etc., on the dead probands. These death certificate data are then verified with respondent(s).
2. *On dead probands and living controls* (information on familial, social and miscellaneous factors including):
   a. Size of sibship.
   b. Birth order.

c. Parental ages at birth of proband.
d. Marital history: number of marriages, divorces, separations, etc.
e. Fertility history: offspring, miscarriages, stillbirths.
f. Occupation.
g. Education.
h. Religion.

3. *On all relatives* (parents, sibs, offspring, spouse, parents and sibs of spouse for both study groups—dead probands and living controls):
   a. For each individual: name, sex, birthplace, year of birth (in addition: birth order for sibs, spouses and spouses' sibs).
   b. For living relatives: age and address (for follow-up checking and possibly longitudinal survivorship study).
   c. For dead relatives: place and year of death, cemetery, respondent's information concerning age at death and cause of death to be checked with death certificates.
   d. Families of dead probands are also checked with respect to consanguinity and for overlapping of family members by use of Living and Dead card files.

B. *Progress of the study*

1. *First phase*
   a. The selection of the study population (tables 2, 3, 4).
   b. Search of death certificates of the 550 selected probands.
   c. Pretest of interview schedule in field trials and final revision of schedule to incorporate all changes deemed feasible on the basis of the pilot study and pretest period.
   d. Design of plan for matching living controls to deceased probands (figures 1–5).
   e. Compilation of an Interviewer's Manual as a standardized guide line for procedures and decisions.
   f. Selection and training of initial staff.
   g. Beginning of a record-keeping system with cross-index files for relatives, as well as probands and controls; and design of various forms to be used in recording data.
   h. Initiation of interviewing of respondents for deceased probands.

2. *Second phase*
   a. (Personnel)
   Recruiting and *training* of qualified personnel (interviewers, neighborhood screeners for matching living controls and clerical staff with replacements as required). The length of training period for individual interviewers depended on their experience, qualifications and aptitude, usually requiring at least 3 to 4 combined group and individual instruction sessions. Screeners received similar, though somewhat less extensive, training than interviewers.
   b. (Procedures and record system)
   Establishment of standardized procedures for each phase of study:
   Sample selection, interviewing, control matching, obtaining, classifying, recording, verifying and coding data.
   c. (Obtaining proband interview data and selection of matching controls)
   Searching for and interviewing respondents for deceased probands:
      (1) As many respondents are interviewed as can be located to obtain complete data indicated in interview schedule. As

of May 31, 1963, interviews concerning 485 of the 550 probands had been at least initiated. Of these, data concerning 274 were completely obtained. Information on an additional 202 was still incomplete and 9 interviews were in return processing (*i.e.*, being edited).

(2) Screening neighborhoods of deceased probands for the purpose of matching living controls to those probands, with additional procedures for "no contact" households.

d. (Status of Control Search)
As of May 31, 1963, the following controls had been located:
426 matched controls
339 true controls
56 secondary controls
31 census-tract matched controls
65 unmatched controls.

e. Interviewing of living controls and/or any additional respondents required for obtaining family, spouse and spouse family data indicated in the interview schedule. The present status of control interviewing is as follows:
319 true and secondary control interviews
201 complete
74 incomplete
44 return processing
23 census-tract matched interviews
26 unmatched control interviews.

f. (Editing, completion of individual records and verification)
(1) Obtaining additional family data by mail and telephone from out of city family members, friends, institutions, etc., as well as local institutions, funeral homes, cemeteries, hospitals, physicians as needed.
(2) Verification of information concerning probands and living and dead family members (living/dead status, age, cause of death, etc.).
Information about dead probands:
(a) Death certificate information is being checked with one or more respondents who are nearest relatives of the deceased proband.
(b) Hospital and other medical records are also to be checked to verify cause of death but this phase has not yet been initiated. Different sources of information, such as Medical Examiner's Reports, etc. are utilized.
(3) Information about relatives of dead probands and living controls:
(a) For dead individuals, death certificates are being obtained to validate data on these deaths.
(b) For living individuals, verification of birth dates and "living status" is being carried out by telephone, by mail and in person. The status is considered verified if the individual himself or a member of his present household or 2 respondents from his side of the family have confirmed it.

g. *Follow up of "no contact" households* and recontact of initial

refusals by mail and in person in order to reduce the number of final refusals and of no contact households.

h. (Coding and data processing)

(1) *Establishment of a code* suitable for punching on IBM cards and programming for electronic computers. (This code not only utilizes individually specific tabulation codes designed for this particular study but also incorporates such standardized forms as U.S. Census Population Occupational Code and Industrial Code and the International Classification of Causes of Death.)

(2) *Coding of completed records* after editing of interviews and verification of information.

(3) *Development of computer programs* for processing of data.

3. *Third phase*

a. Analysis of the collected data

b. Reporting of the results.

Coronary Heart Disease in Relation to Blood Pressure and Cholesterol Levels in Population Studies[1]

By Frederick H. Epstein, M.D.,[2] and Marcus O. Kjelsberg, Ph. D., *Department of Epidemiology and Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, Mich.*

Studies of familial aggregation of disease may be carried out among various subdivisions of the general population. A total community within the general population may be the site of operation. A random sample of a defined population may be drawn. Index cases and their kin may be identified from a number of possible sources and compared with a suitable control group. Each of these approaches carries advantages and shortcomings and there is no single approach which is better than all others, the choice depends on the aims. The subsequent discussion will be mostly confined to familial aggregations of coronary heart disease in the population at large. While previous studies have suggested differences in familial clustering of coronary disease between the relatives of index cases and controls, the picture as it presents itself in the general population is largely unknown. In order to illustrate the community approach, a description of our own investigations will be given and compared with some calculations based on population models.

[1] This study was supported by the Cardiovascular Research Center, University of Michigan, under program project grant H-6378 from the National Heart Institute, National Institutes of Health, U.S. Public Health Service, and prepared in collaboration with the research staff, Cardiovascular Center, University of Michigan, and the Tecumseh Community Health Study.

[2] Supported by a Research Career Award (HE-K6-6748) from the National Heart Institute, National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare.

735-712 O—65——18

SOME OBSERVATIONS ON CORONARY HEART DISEASE IN A TOTAL COMMUNITY

The community which is the site of our studies is the town of Tecumseh, Michigan, some 30 miles from Ann Arbor (Epstein, 1960a; Francis, 1961; Napier, 1962). Within this study area, about 90 percent of all inhabitants have received intensive medical examinations during 1959 and 1960; a second round of examinations is now in progress. The initial examinations included 8,641 persons. In such a setting, the investigation of familial aggregations of disease or, for that matter, health is a natural component of the total operation. Great flexibility is possible in using variously defined case and control groups within this entire universe. The clustering of variables or events within kindreds, families, households, and other population subgroups can be observed against a firm background of age and sex-specific distributions in the total group. A study of this kind in a total, natural community provides the ultimate opportunity to weave man and his environment into an all-embracing picture of human ecology. The advantages of total community studies, as contrasted with studies among population samples, are apparent. Their shortcomings revolve around difficult problems of adequacy in numbers, methods, strategy and logistics. Complexity is now taken for granted in many fields of research. Some problems of population biology require equally complex approaches. There are problems of human ecology which can only be answered within the encompassing framework of an operation of which the Tecumseh Study is an example. The description of such a framework is presented as one approach to the identification of genetic and environmental influences and their effect on the maintenance of health and susceptibility to illness.

One of the first tasks in Tecumseh was the charting of kindreds and the development of a system to assemble individuals related to each other in various ways, so that the attributes of these persons could be compared. Kindred charts were prepared by hand and each member of the kindred was given a 10-digit code which identified his position within the kindred.[3] Households were identified independently. This system makes it possible to analyze the data on living and deceased individuals by biological relatedness, households, and examination status. The population in Tecumseh could be assigned to a total of some 4,500 overlapping kindreds.

Since a major interest centers on coronary heart disease, one of the first questions was whether clinically manifest coronary heart disease could be shown to cluster in families in the Tecumseh population. During the first round of examinations, 248 men and women were identified as having a probable diagnosis of coronary disease by cri-

[3] Samples of kindred charts are obtainable from the authors.

teria serving the needs of this particular analysis. Since most of these persons were middle-aged or older, many of their parents were deceased while many of their children were too young to show clinically detectable disease, and diagnostic methods for pre-clinical disease are, at present, unsatisfactory. Pending more accurate data on causes of death among the parents and more sensitive methods of clinical and pre-clinical disease detection among the children of index cases, a preliminary look at the data was limited, therefore, to the siblings of index cases. A sizeable proportion of the siblings under consideration live outside the study area; in order to obtain additional, direct information, we have interviewed but not yet examined those siblings of index cases who live within a radius of some 30 miles of Tecumseh and have analyzed the information on deaths among the siblings of index cases. The composition of these sibships in terms of examined persons, interviewed persons, siblings living beyond immediate reach and deceased siblings is shown in table 1. Thus, some 54 percent of all living siblings (466 out of 869) have either been examined or live within reach. With regard to the remainder, information already available from the family history must be verified through mail interview and contact with private physicians since it would be beyond our means to send a roving team around the country to examine the 403 siblings involved as well as appropriate controls.

TABLE 1.—*Distribution of brothers and sisters of 248 cases with coronary disease (CHD) in Tecumseh, Michigan, by examination status, diagnosis and location*

Examination status	Diagnosis of CHD	Male N	Female N	Total N	Percent of total
Examined cases	Probable	142	106	248	20.4
Examined sibs	Suspect	7	13	20	
	None	20	24	44	
	Subtotal	27	37	64	5.3
Interviewed sibs (living nearby).	Probable	8	1	9	
	Suspect	5	10	15	
	None	64	66	130	
	Subtotal	77	77	154	12.7
Noninterviewed sibs (not living nearby).		214	189	403	33.2
Sibs who died below age 30	Death due to: Heart Disease	2	2	4	
	Other Causes	71	64	135	
Sibs who died at age 30 or over	Death due to: Heart Disease	34	22	56	
	Other Causes	75	75	150	
	Subtotal	182	163	345	28.4
All sibs		642	572	1,214	100.0

These data have been reviewed in detail because they raise a fundamental problem in the study of familial aggregations among general populations in the United States. The mobility of the population in this country makes it difficult to assemble complete and accurate information on kindreds. The distribution presented gives a quantitative assessment of this problem in a general population.

Multiple cases among examined persons were observed in six sibships so that, in 236 of the sibships, the index case was the only examined individual with coronary disease in the sibship whereas the

remaining 12 index cases occurred in pairs in the remaining six sibships. In addition, there were 10 sibships in which an interviewed brother or sister of an index case also reported "probable" coronary heart disease; in one of these 10 sibships, two other brothers had died of "heart trouble". There were another 43 sibships in which a sibling of an index case reportedly died of heart disease. Thus, there remained 133 sibships of index cases in which there was neither another living sibling with "probable" coronary heart disease among persons examined or interviewed nor deaths from this disease as far as could be ascertained.

It may be asked to what extent incompleteness of kindred data might limit the inferences which can be drawn. If, among all living family members, those available were a random sample of the total group, the analysis could be confined to those within reach. In fact, the available family members may well be different since, in some studies, the more mobile segments of the population tend to have more disease (Syme *et. al.*, 1964). Whether these persons differ also in their genetic constitution is, however, another question which would be difficult but possible to test. As far as the feasibility of carrying out familial studies of chronic disease in the general population in the United States is concerned, one has to face the facts and develop statistically valid epidemiological methods to solve the problem as it presents itself in a country such as this. This appears preferable to taking a nihilistic approach and deciding that such studies can only be done in less mobile populations. Since the disorders under discussion no doubt result from both genetic and environmental influences, the etiology of these conditions must be studied as part of the environment in which they occur and observed in different geographic locations rather than through extrapolation of findings from one area to another.

AGGREGATIONS OF CORONARY HEART DISEASE IN FAMILIES

As a first approach, it might be useful to view familial aggregations of coronary heart disease in terms of a hypothetical model, for comparison with the situation encountered in the field. A disease as common as coronary heart disease must inevitably coexist sometimes in several members of the same family by chance alone. This feature, as is well known, causes difficulty when familial aggregations are studied by *sampling* index cases and controls, the "fallacy of the index case" phenomenon; this particular problem does not exist in studies among total populations. Let us assume that all middle-aged men in a population are brother-pairs (table 2). In a population of 2,000 men, there would, therefore, be 1,000 brother-pairs.

The prevalence of manifest coronary heart disease among middle-aged men in this country is approximately 5.5 percent (Epstein, 1960b) so that there will be 110 cases and 6 of these will be expected to occur in 3 pairs (0.055 times 0.055 being 0.0030). Thus, if there were no familial aggregations of coronary heart disease, an epidemiological study of 1,000 living middle-aged men and their brothers would yield 3 instances in which both brothers are affected simultaneously. At the present state of knowledge, it is not possible to say to what extent the true coexistence of disease among living male siblings might exceed chance expectation. However, even if the frequency of concordant pairs were four times greater than expected, a sizable excess, no more than 12 out of 1,000 sets of brother-pairs would be affected simultaneously. The data may be viewed in another way. Thus, with four times random expectation, 24 out of the total of 110 cases, *i.e.*, a little more than 20 percent, occur in pairs, as compared with a random expectation of about 5.5 percent (6 out of 110).

TABLE 2.—*Frequency of coronary heart disease in a hypothetical population of 1,000 brother-pairs*

Number of examined members per sibship	Total number of persons in sibships	Number of cases per sibship	Total number of sibships		Total number of cases	
			Random expectation	4 times random expectation	Random expectation	4 times random expectation
2	2,000 (1,000 sibships).	2	3	12	6	24
		1	104	86	104	86
		0	893	902	0	0
Total	2,000		1,000	1,000	110	110

Frequencies are based on a prevalence of 5.5 percent for coronary heart disease among middle-aged men.

Detecting familial aggregations would be easier if the disease prevalence were greater. A prevalence of around 5 percent encompasses only those cases of coronary heart disease which are clinically detectable by presently available means. The true prevalence of severe lesions, based on autopsy findings, is in the region of 20 percent. In this situation, four times random expectation (40 concordant brother-pairs) would yield 160 brother-pairs simultaneously affected among a total population of 2,000 brothers.

Against this background, some very preliminary calculations based on the Tecumseh data may be discussed. We are not yet ready to present definitive data on familial aggregations of coronary disease in Tecumseh, but merely intend to discuss ways to approach the problems of analysis. We wished to calculate whether myocardial infarction occurred more often than could be expected by chance in the sibships of persons examined in Tecumseh. This pilot analysis was limited to sibships in which all examined members were aged 40 years or older. In order to increase the yield, persons with a history of both

"probable" and "suspect" myocardial infarction were combined into a single category, "myocardial infarction". By contrast, the term "coronary heart disease", as used earlier, included persons who had either a history of probable myocardial infarction, symptoms of probable angina pectoris, electrocardiographic evidence of myocardial damage, or a combination of these. Each man and woman in the sibships under consideration was assigned a probability of having myocardial infarction, as defined. The probabilities assigned were simply age- and sex-specific prevalence rates for probable or suspect myocardial infarcation in the total Tecumseh population. A computer program was developed to calculate the number of sibships, according to the number of examined persons in each sibship, in which multiple cases would be expected to coexist by chance alone. These expected numbers were compared with the actual numbers observed (table 3).

TABLE 3.—*Frequency of observed and expected* number of sibships** and cases with a diagnosis of myocardial infarction (probable and suspect); Tecumseh, Michigan, 1959–1960*

Number of examined members per sibship	Total number of examined persons in sibship	Number of cases per sibship	Total number of sibships		Total number of cases	
			Observed	Expected	Observed	Expected
2	264 (132 sibships)	2	3	.8	6	1.6
		1	18	17.1	18	17.1
		0	111	114.1	0	.0
3	81 (27 sibships)	3	0	.0	0	.0
		2	0	.5	0	1.0
		1	1	5.1	1	5.1
		0	26	21.4	0	.0
4	16 (4 sibships)	4	0	.0	0	.0
		3	0	.0	0	.0
		2	0	.2	0	.4
		1	0	.9	0	.9
		0	4	2.9	0	.0
Total	361 (163 sibships)		163	163.0	25	25.7

*Expected frequencies based on age-sex-specific prevalence rates for Tecumseh population.
**Including only those sibships with all examined members age 40 or over.

There were 361 men and women in 163 sibships including 2, 3 or 4 examined members. Of the two-member sibships, 0.8 sibships were expected to show both members to be diseased against an observed number of three sibships in which both members had a diagnosis of probable or suspect myocardial infarction. The remaining observed and expected numbers agreed quite closely. Viewed differently, it may be seen that 6 of the observed 25 cases were, in fact, multiple cases within the same sibships, as compared with an expected frequency of 1.6 cases; whereas the remaining cases occurred singly in the sibships of various sizes. It may be noted that the prevalence of probable and suspect myocardial infarcation in this particular group is about 7 percent (25 cases in 361 persons), as compared with a prevalence of 5.5 percent in the hypothetical model discussed before.

The lesson learned up to this point, using either the hypothetical model or the preliminary test in Tecumseh, suggests that there is a significant problem relating to the adequacy of numbers in searching for familial aggregations of coronary heart disease in the general population. This problem exists not only in the case of coronary heart disease but also for other chronic diseases. Actually, familial aggregations of coronary disease are probably more marked than these calculations would suggest. It is likely that such aggregations could be more readily demonstrated if information on all family members were more accurate and complete. Thus, how can one be sure that a person who is labelled as being free from disease might not have severe coronary atherosclerosis, diagnostic methods being so insensitive? In a person who has died, how sure can one be of the accuracy of a diagnosis of arteriosclerotic heart disease as a cause of death and how many such cases may have been misclassified under another label (Beadenkopf *et. al.*, 1963)? Another uncertainty relates to the difficulties in obtaining accurate information on persons who cannot be directly interviewed and examined. A further problem arises because the frequency of the disease is highly dependent on age, so that freedom from manifest disease at a younger age would not have the same significance as the same finding in an older person of the same sex. This makes it difficult to compare individuals and groups differing in age. Similarly, freedom from disease at a given age has not the same significance in men and women. Therefore, sibship comparisons should be age- and sex-specific, is—difficult to accomplish on account of the problem of numbers; alternatively expected and observed rates, adjusted for age and sex, may be contrasted, as illustrated earlier. There is also a question whether data on living and deceased persons can be combined in an analysis as has been done in some studies. Thus, would it be reasonable to say that a man whose brother or father has just died of a myocardial infarction at the age of 75 or 80 has more of a "family history" of the disease than another man whose brother or father is well and alive at that age? In other words, for a disease which is almost by necessity the major cause of death in older people, is it reasonable to give the same weight to these deaths, regardless of the age at which they occur? It is not possible, therefore, to combine data on living and deceased persons without giving special consideration to the problem of age at death. At the same time, data on living family members should not be considered without taking into account persons in the kindred who have died. Failure to detect familial aggregations of coronary disease among living members might be due, for instance, to the fact that relatives with the disease have died prior to the study. It is clear, therefore, that analysis should not be confined solely to the survivors. An attempt to take into consideration deceased relatives

has been made in the analysis of the data from the Tecumseh study, discussed earlier.

The nature of these problems makes it unlikely that there is a single approach to their solution. In the field, a number of different ways of analyzing the data must be used to gain a composite picture from several pieces of evidence. One approach has already been illustrated in terms of comparing observed and expected frequencies. Other approaches might be briefly considered.

1. The mortality experience of the parents or siblings of propositi and controls may be compared. The proportion of parents or siblings living and deceased and the frequency distribution of the ages at death may be ascertained for the two groups. Life tables on cohorts might be constructed for the parents or siblings in the two groups and the survivorship curves contrasted. Such an analysis might suggest that changing environmental factors affect the relatives of propositi differently from those of controls.

2. In a study like the one in Tecumseh, one can compare morbidity and mortality among relatives living outside the study area. A greater frequency of disease among the relatives of index cases might indicate that one or more familial factors are operative.

3. As a further test for familial aggregations, the frequency of "suspect" disease among relatives of propositi and controls can be estimated.

We will report on some of these approaches in the Tecumseh population when a more complete report can be given. In the meantime, it must be stressed that the problems discussed so far relate to the situation where data on a population include only currently present conditions and retrospective information. Longitudinal, prospective observations, as will be mentioned, will add greatly to the potential of such studies.

PREDISPOSING FACTORS AND FAMILIAL AGGREGATIONS

Up to this point, attention has been confined to the question whether coronary heart disease aggregates in families, regardless of the mechanisms involved in causing these aggregations. In fact, on *a priori* grounds, there can be no doubt that coronary disease must show familial clustering since the disease is clearly associated, in individuals, with three variables which are, in part, determined by heredity: serum cholesterol, blood pressure and glucose tolerance, using these terms in preference to hypercholesterolemia, hypertension and diabetes. In view of these relationships, it is rather surprising that familial aggregations of coronary disease are, as was shown, so difficult to demonstrate in a general population. A recent review of

familial studies of coronary heart disease has pointed out the shortcomings of existing data and the need for new studies, using both the index case-control and community type of approach (Epstein, 1964). The subsequent discussion will be confined to serum cholesterol and blood pressure levels even though there is a great need to search for possible genetic influences which may exist independently of these two recognized associations.

In familial studies, it would facilitate some types of analysis if persons could be properly subdivided into those with and without hypercholesterolemia or hypertension. The data from Tecumseh support the view that such subdivisions are artificial (Johnson *et al.*, 1965). While the total evidence tends to favor the opinion that both serum cholesterol and blood pressure levels are determined by multiple genes and multiple environmental influences, it is convenient to divide the distributions arbitrarily into an upper and lower range by suitable cutting points, in order to estimate the effect of these variables on the occurrence of coronary heart disease. Toward this end, some further calculations based on the hypothetical population model of 1,000 brother-pairs used earlier are instructive. It will be assumed, for the purpose of these calculations, that hypercholesterolemia and hypertension are both inherited as dominant characters. The probable incorrectness of this assumption has already been pointed out but, from a practical point of view, both of these variables can be considered to behave *as if* their inheritance were dominant, using a suitable cutting point between "normal" and "abnormal." This was illustrated by Platt (Platt, 1959), using the data of Pickering and his associates (Hamilton *et al.*, 1960) on blood pressures which are, in the opinion of most workers, unimodally distributed. At face value, even unimodally distributed data can be manipulated in such a way that the result is compatible with a dominant mode of inheritance, as discussed and criticized earlier by Oldham and associates (Oldham *et al.*, 1960). For the purpose of the calculations which follow, it will be assumed, therefore, that blood pressure and also serum cholesterol levels, which show a similar familial pattern, are distributed in the general population *as if* each were determined by a single, dominant gene; whether this assumption is correct or not has no influence on the broad interpretation of the subsequent calculations. It is further understood that hypercholesterolemia and hypertension do not appear to be significantly associated (Dawber *et al.*, 1957).

Before turning to the concrete situation, regarding the interrelationship between coronary disease and elevated serum cholesterol and blood pressure levels, let us assume that the population can be subdivided into persons with a factor, to be designated as X, inherited as a dominant character, which carries an increased risk toward the development of coronary disease, and persons without this predisposing factor, to be designated as x. Knowing the frequency of X in the

population of 1,000 brother-pairs, it can be calculated under the assumption of random mating by the use of the appropriate formulae (table 4) how many brother-pairs are characterized by the combinations X-X, X-x, and x-x. Taking the frequency of X as 30 percent, these pairs will occur with a frequency of 190, 219 and 591, respectively, among the 1,000 brother-pairs (table 4). Next, we will again set the prevalence of coronary heart disease among these men at 5.5 percent and assume that the disease is three times as frequent among men with X as among those with x. On these premises, one can calculate from the formulae in table 4 the distribution of cases of coronary disease in the three types of sibships X-X, X-x and x-x according to whether the brothers are concordant or discordant for X and x. It is found that 3.5 (2.0+0.8+0.7) of the 1,000 sibships will show both brothers affected, representing 7 cases (4.0+1.6+1.4) among the total of 110 cases (table 4). It will be recalled that the random expectation of simultaneously affected pairs was 3 per thousand. Surprisingly, no more than an additional 0.5 pairs with both brothers affected were added above random expectation.

TABLE 4.—*Frequency of coronary heart disease in a hypothetical population of 1,000 brother-pairs, according to presence or absence of a characteristic X*

Type of sibship	Total number of sibships	Number of cases of CHD per sibship	Total number of sibships	Total number of cases
X-X	$N_1=N(p/4)(4+4pq+pq^2)$ =190 (380 persons)	2	$k^2R^2N_1=2.0$	4.0
		1	$2kR(1-kR)N_1=35.0$	35.0
		0	$(1-kR)^2N_1=152.5$	
X-x	$N_2=N(pq^2/4)(6+2q)$ =219 (438 persons)	2	$kR^2N_2=0.8$	1.6
		1	$[R(1-kR)+kR(1-R)]N_2=28.5$	28.5
		0	$(1-kR)(1-R)N_2=190.0$	
x-x	$N_3=N(q^2/4)(1+q)^2$ =591 (1,182 persons)	2	$R^2N_3=0.7$	1.4
		1	$2R(1-R)N_3=39.5$	39.5
		0	$(1-R)^2N_3=551.0$	
Total	$N=N_1+N_2+N_3=1,000$ (2,000 persons)		$N=N_1+N_2+N_3=1,000$	110.0

Frequencies are based on a prevalence of 5.5 percent for coronary heart disease (CHD) among middle-aged men.
Calculations assume dominant inheritance for X and a frequency for X of 30 percent.
X-X indicates that both brothers of the pair have characteristic X; X-x indicates that one brother of the pair has characteristic X; x-x indicates that neither brother of the pair has characteristic X.
q^2=proportion of x individuals in the population.
$p=1-q$.
R=prevalence rate for CHD among the x individuals.
kR=prevalence rate for CHD among the X individuals.
The numerical values used were $N=1,000$, $q^2=.70$, $R=.0344$ and $k=3$.

With some hesitancy, let us consider that X stands for the combined frequency of hypercholesterolemia and hypertension[4] in a population of middle-aged men. In fact, in such a population in this country, these conditions do occur with an approximate frequency of 30 percent, as assumed in the foregoing calculations, this figure being slightly lower than found in Framingham (Dawber *et al.*, 1957) but slightly

[4] Serum cholesterol 260 mg. percent or over; blood pressure 160 mm Hg. systolic or over and/or 96 mm Hg. diastolic or over.

higher than the Tecumseh experience. On the basis of this highly simplified hypothetical model, therefore, we have attempted to explain why familial clustering of coronary disease—and this applies to other chronic diseases as well—is difficult to demonstrate in a general population in a one-time, cross sectional study. Are these statements in contrast to the relatively few studies in which relatives of index cases showed, perhaps, a three- or four-fold excess of coronary disease as compared with relatives of control subjects (Epstein, 1964)? In view of the fact that these data include or are based on information on deceased family members, and on histories obtained from family members rather than direct examinations, there is no need to postulate that there is necessarily any contradiction.

A final look at the population model reveals that 69.1 (4.0+35.0+1.6+28.5) of the 110 cases, or as many as 63 percent occurred among men who might be considered as belonging to hypercholesterolemic or hypertensive families. In the Tecumseh population, in the five out of the six sibships with multiple examined cases of coronary disease, there was at least one person with at least one of these variables in the upper range. Moreover, among 81 men with coronary disease aged 40–54 in Tecumseh, 65 percent showed serum cholesterol, blood pressure or both in the upper range, the corresponding figure among 52 women being 75 percent. It is suggested that the factors, both genetic and environmental, which contribute to elevated serum cholesterol and blood pressure levels, account in considerable part for aggregations of coronary heart disease in families.

It is of interest to look at the population model also from the point of view of the investigator who would sample index cases and controls at random from this universe of 2,000 middle-aged brothers. In a 10 percent random sample of index cases and controls, 0.7 cases of coronary disease would be expected among the 11 siblings of index cases—a prevalence of 6.4 percent, as compared with a prevalence of 3.6 percent among 186 siblings of controls. This difference, though less than twofold, is nevertheless not negligible.

OUTLOOK AND CONCLUSIONS

In the foregoing discussions, an attempt was made to show that the search for familial aggregations of coronary heart disease and other chronic disorders of similar frequency of occurrence in the general population, on the basis of a *single* survey, will yield relatively few instances of multiple cases within sibships or, presumably, kindreds. Awareness of this problem is essential in planning such studies. Fortunately, seen in perspective, the problem has a solution. First of all, in a prospective study of coronary heart disease, new cases will accumulate rapidly. Since the prevalence of the disease among

middle-aged men in this country is about 5 percent and the yearly incidence about 1.5 percent (Epstein, 1960b), after an observation period of 3 or 4 years as many new cases will accumulate as were identified at the beginning. Thus, with each year of observation, the chances increase to detect the occurrence of new events where they would be expected to arise preferentially—*i.e.*, among the relatives of the index cases identified at the outset. Secondly, the primary interest relates not to the detection of established disease within families but to the familial clustering of predisposing factors since, from the overriding point of view of disease prevention, it is the early detection of susceptibility to the future development of overt illness which commands major attention. The prevalence of these predisposing factors, such as elevated serum cholesterol and blood pressure levels as usually defined, is considerably greater than the frequency of the associated disease itself. The problem of relatively small numbers is thus reduced. There is much hope, therefore, that the extent of familial clustering of coronary heart disease and predisposing factors in the population at large will become known within the next few years.

SUMMARY

Aggregations of coronary heart disease in families are difficult to demonstrate in one-time, cross-sectional epidemiological studies of total populations. The index case-control type of approach may be misleading in this regard unless index cases and controls are representative of the universe from which they are derived. An actual population study and a population model have been used to illustrate and discuss the methodological problems involved in the study of familial clustering of coronary heart disease and, by implication, other chronic disorders. Even though two major factors predisposing to coronary atherosclerosis, elevated serum cholesterol and blood pressure levels, show definite familial trends, their effect on familial concentrations of coronary disease was shown to be surprisingly small on the basis of some theoretical calculations. It is suggested that prospective rather than cross-sectional studies are needed to assess the true degree of familial aggregations of coronary heart disease.

REFERENCES

1. Beadenkopf, W. G., Abrahams, M., Daoud, A. and Marks, R. U., 1963. An Assessment of Certain Medical Aspects of Death Certificate Data for Epidemiologic Study of Arteriosclerotic Heart Disease. J. Chronic Dis. 16: 249–262.
2. Dawber, T. R., Moore, F. E. and Mann, G. V., 1957. Coronary Heart Disease in the Framingham Study. Am. J. Pub. Health 47. Supplement to No. 4: p. 4–24.

3. Epstein, F. H., 1960a. An Epidemiological Study in a Total Community: The Tecumseh Project. University of Michigan Medical Bulletin 26: 307–314.
4. Epstein, F. H., 1960b. Epidemiology of Coronary Heart Disease. In: Modern Trends in Cardiology, edited by A. Morgan Jones. London, 1960. Butterworth.
5. Epstein, F. H., 1964. Hereditary Aspects of Coronary Heart Disease. Am. Heart J. 67: 445–456.
6. Francis, T., Jr., 1961. Aspects of the Tecumseh Study. Public Health Reports 76: 963–965.
7. Hamilton, M., Pickering, G. W., Fraser Roberts, J. A., and Sowry, G. S. C., 1954. The Etiology of Essential Hypertension. IV. The Role of Inheritance. Clin. Sci. 13: 273–303.
8. Johnson, B. C., Epstein, F. H., Kjelsberg, M. O., 1965. Distributions and Familial Studies of Blood Pressure and Serum Cholesterol Levels in a Total Community—Tecumseh, Michigan. J. Chronic Dis. (In Press).
9. Napier, J. A., 1962. Field Methods and Responses Rates in the Tecumseh Community Health Study. American Journal of Public Health 52: 208–216.
10. Oldham, P. D., Pickering, G., Fraser Roberts, J. A., and Sowry, G. S. C., 1960. The Nature of Essential Hypertension. Lancet 1: 1085–1093.
11. Platt, R., 1959. The Nature of Essential Hypertension. Lancet 2: 55–57.
12. Syme, L., Hyman, M., and Enterline, P. E., 1964. Some Social and Cultural Factors Associated with the Occurrence of Coronary Heart Disease. J. Chronic Dis. 17: 277–289.

Inherited Antigenic Differences in Serum Beta Lipoproteins in Relation to Coronary Artery Disease and Diabetes

By Baruch S. Blumberg, M.D., Barbara B. Mishell, B.A., and Sam Visnich, B.S., *National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare, Bethesda, Md.*

INTRODUCTION

The method of genetic analysis described in this paper originates from the age-old observation that people differ in their susceptibility to illness. All those exposed to the tuberculosis organism do not develop clinical disease nor does everyone in an environment where rheumatic fever is common fall prey to it. The same may be said for a large number of infectious and noninfectious diseases. Barring those cases in which chance or divine intercession are operating, this difference in susceptibility must be due to inherent differences between individuals, and some of these differences are inherited. The study of normal inherited variation may, therefore, help in differentiating those human differences which bear on disease resistance and susceptibility. Biochemical polymorphic traits are particularly useful for this purpose. In polymorphisms two or more discontinuous forms of an inherited trait are common in a population, and the frequency of the least common form is higher than could be expected from recurrent mutation alone (Ford, 1956). It has been theorized that a polymorphism is maintained as a result of differences in selective value of the genotypes; and that the heterozygote is at a selective advantage as compared to either of the alternate genotypes. Selection is expressed as differential fertility, that is, the relative contribution of the different genotypes to the genetic makeup of the next reproducing generation. Selection can take place at any time in the reproductive cycle or before the completion of childbearing, and any diseases which

have a differential selective effect on the genotypes could influence the gene balance. A list of polymorphic traits which have been studied in human populations is given in table 1 (Blumberg, 1961).

TABLE 1.—*Some polymorphic traits in humans*

Red blood cell types	Serum proteins
ABO MNS P Rhesus Lutheran Kell Lewis Duffy Kidd Diego Sutter Sex-linked	Haptoglobins. Transferrin. Gamma globulins. Beta-lipoprotein. Group-specific substance. Serum cholinesterase.
Other cell types	**Miscellaneous**
Hemoglobin G6PD White blood cell platelet	BAIB, urinary excretion. Secretor of ABH blood group substance. Taste of PTC.

GENERAL OUTLINE OF THE METHOD OF INVESTIGATION

1. *The detection of discontinuous, biochemical differences between individuals.*

"Discontinuous" indicates that the qualitatively different forms of a trait are readily distinguishable from each other. This makes it possible to clearly designate an individual as of one or another type, and ambiguities in classification would occur only rarely. In continuous traits, the differences are quantitative and classification is dependent on arbitrary divisions of the spectrum. In some cases, there can be two curves of continuous distribution with a definite division between them, and for practical purposes the trait can be considered as discontinuous.

Many discontinuous biochemical differences between individuals have been found by the application of protein separation and identification techniques to nonpooled blood samples. When Smithies (1955) developed the technique of starch gel electrophoresis, he was able to separate serum protein into more fractions than had been possible with previous methods. He noted that the protein patterns were not the same in all individuals; this led to the discovery of the haptoglobin polymorphism (Smithies and Connell, 1959), and later the transferrin polymorphism (Smithies and Hiller, 1959), the post albumin polymorphism (Smithies, 1959), the thyroxine-binding prealbumin differences in *Macaca mulatta* monkeys (Blumberg and Robbins, 1960), and others. Similarly, the development of rapid and

sensitive methods of immunoelectrophoresis (Grabar and Williams, 1955) led to the discovery of the Gc group-specific serum system (Hirschfeld and Beckman, 1960). The discovery of discontinuous differences is not limited to blood studies. Electrophoretic differences observed in cattle milk proteins resulted in the identification of the β-lactoglobulin polymorphism (Aschaffenburg and Drewry, 1955) and later in Fulani and Indian cattle, the alpha-lactalbumin system (Blumberg and Tombs, 1958). Studies on urine have revealed the polymorphism for the excretion of the amino acid, beta-amino isobutyric acid (BAIB) (Gartler, 1959).

One of the most rewarding means for detecting discontinuous biochemical differences is by the study of inherited differences in sensitivity to drugs, that is, by pharmacogenetics. The pseudocholinesterase polymorphism (Kalow, 1959) and the very important G6PD polymorphism were discovered in this manner (Hockwald *et al.*, 1952; Beutler, 1959).

2. *To determine if the discontinuous differences are inherited.*

a) *Is the trait an inherent characteristic of an individual?*

In order to determine if an individual's phenotype is always the same, or, if not, under what conditions it might vary, the phenotype is determined over the course of several days, weeks, or years, and if possible, under varying conditions of nutrition, climate, and any other factor which could conceivably alter it. For example, an individual's erythrocyte sedimentation rate may vary depending on his clinical condition; it could not be considered an unvarying characteristic of the person by which he could be identified at any time. His hemoglobin type (*i.e.*, A/A, A/S, S/S) would nearly always be the same and detectable, barring some extraordinary pathological change.

b) *Age dependence.*

The late development of an inherited trait often makes genetic analysis difficult. A well known example in medical genetics is Huntington's chorea, an inherited disease which may become apparent only late in life. Pedigree analysis of families with many young children may show a marked difference from the predicted frequencies, and from the frequencies determined in families in which the offspring have all reached the age at which the disease can develop.

Age dependence is usually not a difficult problem for the analysis of biochemical polymorphic traits, although caution must be observed. The haptoglobins of many infants are not fully developed until approximately 4 months of age and reliable phenotyping cannot be done before that time. A child's Gm (gamma globulin) type is not expressed until sometime after birth; furthermore, the mother's specificity may be transmitted transplacentally and detected in the child's blood. There are, however, some instances when the age dependence may extend over a longer period of time. Steinberg and Giles (1959) have

735-712 O—65——19

described an inherited polymorphic trait with a striking age dependence.

c) *Sex distribution.*

Most of the polymorphisms are autosomal and independent of sex. However, the sex dependence has not been carefully studied, even for some of the better understood traits. At least two sex-linked biochemical polymorphisms are now known: the G6PD deficiencies and the Xg^a blood group systems. For a newly discovered discontinuous trait, sex dependence can be determined by appropriate population and family investigations.

d) *Twin studies.*

Twin studies are very useful in determining if a discontinuous trait is inherited, although they do not give very much information on the mode of inheritance. If all identical twins are concordant, if non-identical twins are either concordant or discordant, and if the difference is significant, then the likelihood that the discontinuous trait is inherited is very high. It is particularly useful if the trait can be studied on stored material, such as serum, for then a single collection of blood samples from a twin population can be used to study a variety of traits. Sera collected from twins for one population study (McDonough *et al.*, 1963) have been retained in our laboratory for several years and used for subsequent investigation of serum polymorphisms. Twin studies are a rigorous test of a genetic hypothesis. If even one of the proven identical twin pairs are discordant for the trait, this will cast doubt on a simple genetic hypothesis or, depending on how strongly held the convictions are, on the monozygosity of the twins.

e) *Family studies.*

To learn the mode of inheritance, family and population studies are necessary. For many biochemical traits it is not possible to determine if a particular family is useful for genetic testing until the biochemical study is actually performed. However, for polymorphic traits, if two or more of the forms are relatively common and particularly if the heterozygote can be detected, then all families studied contribute to the genetic analysis. The method of analysis proposed by Smith (1956) is particularly helpful in efficiently utilizing the information, and this has been used in several studies.

For many of the traits, it is possible to use stored material. In our laboratory a file of family sera is maintained for use in the appropriate investigations.

f) *Racial differences.*

The frequency of the genes controlling polymorphic traits may vary enormously in different populations. The extreme of this is seen in traits which are essentially absent in some populations, but polymorphic in others: *i.e.*, the Diego blood group is relatively common in some Asian and American Indian populations, but extremely rare in populations of European and African origin (Levine *et al.*, 1956);

the Js red blood cell group is common in Africans and populations of African origin, but rare or absent in other populations (Giblett, 1958).

For many of the biochemical traits, it isn't possible to ascertain the phenotype without actually performing the biochemical test. Hence, it is difficult to know which families to study to obtain critical matings for testing the genetic hypothesis. The usual practice is to collect as many families as possible in a given population, determine the phenotypes of each entire family, note the pattern of segregation, and from this propose and test a genetic hypothesis. In order to decide in which population to conduct the family (and twin) studies, the phenotype frequencies in several populations are determined. (For this purpose, collections of sera from different populations are extremely useful; some of the material available in our own laboratory is shown in table 2.) With this information, an intelligent selection of the appropriate population for family studies can be made. For example, if there are two forms of a trait, A_1 and A_2, and in population I, A_1 is 99 percent and A_2 1 percent, but in population II, A_1 is 50 percent and A_2 50 percent, then population II would be far more convenient for the study of the genetics of the trait.

TABLE 2.—*Populations represented in serum bank*

America:
- Eskimo:
  - Eskimos, Alaska
- Indian:
  - Athabascans, Alaska
  - Athabascans, Canada
  - Haida, Canada
  - Montagnais, Labrador
  - Naskapi, Labrador
  - Quechua, Peru
  - Sioux, South Dakota
  - Tlingit, Alaska
- Negro:
  - Evans County, Georgia
  - Sapelo Island, Georgia
  - Washington, D.C.
- White:
  - Bethesda, Maryland
  - Boston, Massachusetts
  - Evans County, Georgia
  - New York, New York
  - Tangier Island, Maryland

Africa:
- Fulani, Nigeria
- Yoruba, Nigeria

Asia:
- Arabs, Jordan
- Chinese, Formosa
- Chinese, San Francisco
- Israelis, Israel
- Japanese, San Francisco
- Vietnamese, South Vietnam

Pacific:
- Aborigines, Australia
- Micronesians, Majuro
- Micronesians, Rongelap
- Micronesians, Utirik
- Polynesians, Bora Bora

3. *Biochemical nature of the inherited trait.*

An advantage of studying biochemical polymorphic traits is that the investigator is supplied with some "hints" as to the physiological or pathological role of the polymorphism, while this may not be the case when investigating morphological traits. For example, when Smithies determined that the haptoglobin types were inherited, a large body of information on the clinical, pathological, and physiological characteristics of the protein was available to provide a starting point for investigation. Jayle and his colleagues (Jayle *et al.*, 1952) had for years been studying the variation in the amount of this protein in a variety of illnesses, and it was known that the haptoglobins were responsive to inflammatory diseases. It was also known that haptoglobin combined with hemoglobin, and that this might be one of its physiological functions. This led to investigations of the role

of haptoglobin in hemolytic diseases such as the hemoglobinurias, malaria, thalassemia, sickle cell disease, and others (for discussion see Blumberg *et al.*, 1963). Similarly, the role of the transferrins or siderophilins in iron binding was known before the discovery that these proteins were polymorphic. This gave an immediate clue as to the physiological function which might be first studied. There is, however, no good evidence at present that the iron-binding function is implicated in the selective forces which control this complex inherited trait.

A further application of biochemical polymorphisms is in the study of the genetic control of protein synthesis. They provide readily identifiable common traits which can be used in combined biochemical and genetic studies. The brilliant investigations of Smithies and his colleagues (1962) and Parker and Bearn (1963) on the genetic control of haptoglobin synthesis are excellent examples. Polymorphisms also provide suitable systems for the study of population genetics of human cells in which these biochemical traits can be used as markers. The studies of Gartler *et al.* (1962) and others, using the G6PD trait, are examples of the application of this method.

4. *Relation of the polymorphic traits to disease.*

a) *Direct association of a genotype with disease.*

In some instances one of the genotypes may be directly associated with a disease entity. An example of this is the sickle cell homozygote. In some polymorphic systems there is no apparent direct association with illness, but when individuals with one or more of the genotypes are exposed to special hazards, symptoms may occur. Under normal conditions, bearers of the G6PD deficiency appear to be quite normal. However, if they are exposed to a special hazard, such as eating fava beans, or taking any of a large number of drugs, they can develop hemolytic anemia. Sickle cell heterozygotes appear to be clinically normal, but when exposed to a special hazard such as the low pressures experienced in high-flying nonpressurized aircraft, they are more likely to suffer splenic infarct than individuals of different genotype. Similarly, individuals with the rare pseudocholinesterase variant are under ordinary circumstances quite normal; however, if they are given certain muscle relaxants, such as suxamethonium, a prolonged apnea will result.

This group of polymorphic traits are of particular interest because they lend themselves to practical use in preventive medicine. By means of simple chemical tests, individuals who have an increased inherited susceptibility to known environmental hazards can be detected and steps taken to avoid exposure.

b) *Association of rare alleles or phenotypes with an illness.*

Another point of contact of polymorphic systems and clinical disease is the association of a rare allele or phenotype with an illness. There are few well-documented cases of this kind. Ahaptoglobinemia,

which is rare in most non-African populations, may be associated with a variety of hemolytic illnesses, as noted above.

Many rare alleles do not appear to be associated with illness, however. For example, most of the rare transferrin phenotypes have been reported in apparently normal persons (see, for example, Robinson *et al.*, 1963).

c) *Statistical association of polymorphic traits with disease.*

For most polymorphic systems the disease association is not as direct as those described above. However, for many there is a statistical association which in some cases may be quite striking. The systems which have been most studied in this regard are the blood groups. These have been reviewed by Roberts in this symposium and will not be discussed here.

AN EXAMPLE OF A POLYMORPHISM STUDY: THE Ag SYSTEM, AN INHERITED DIFFERENCE IN THE β-LIPOPROTEIN

The first part of this paper was devoted to a general discussion of how biochemical polymorphic traits are detected and analyzed in respect to their genetics, biochemistry, and disease association. In the present section, these methods will be illustrated by the use of a polymorphic trait that has been studied in our laboratory.

The Ag system was discovered as a result of a search for a polymorphism rather than as a chance observation. Experience over the past 10 years has shown that there are large numbers of serum protein polymorphic systems. These include the haptoglobins, transferrins, Gm gamma globulin groups, Gc group-specific alpha$_2$ globulin system, and others. The frequencies of the genes controlling some of these systems are such that if a patient were to receive even a small number of transfusions, the probability of transfusion with a serum protein slightly different from his own would be very high. If this were the case, then the patient might develop an antibody against the foreign protein.

The method of double diffusion in agar gels was used in the investigation. The serum of a transfused patient was placed in the center of a 7-hole Ouchterlony pattern. The sera of normal individuals who could have been, but were not necessarily donors for the patient, were placed in the peripheral wells. After approximately 15 patients had been studied, one was found who gave a strong precipitin line with some but not all the donor sera (Blumberg, 1961). It was immediately apparent that this antiserum could serve as a reagent to study what appeared to be a new polymorphic system. Biochemical and immunological studies showed that the antibody present in the patient's serum was a 7S gamma globulin and that the antigen in the normal serum with which it reacted was the low-density beta-lipoprotein (Blumberg, Dray, and Robinson, 1962). The ability of a serum to react was found to be an inherent characteristic of an individual.

Sera obtained from subjects over the course of several years and under different conditions were tested and found always to give consistent results (Allison and Blumberg, 1961; Blumberg, Bernanke; and Allison, 1962). Furthermore, serum which had been frozen and stored for up to four years still retained the ability to react with the antiserum. This permitted the use of stored serum from our serum bank for many of these studies. Twin studies indicated that it was highly probable that the reactor specificity was inherited (Blumberg, Bernanke; and Allison, 1962), and the family studies were consistent with the hypothesis that the reactor specificity was inherited as a simple autosomal dominant trait (Blumberg, Bernanke; and Allison, 1962; Blumberg, 1963). The reactor phenotype was called Ag(a+), the nonreactor phenotype Ag(a−), the gene controlling the reactor phenotype Ag^A, and the alternate gene whose product had not been discovered *Ag*. Individuals homozygous or heterozygous for the dominant gene Ag^A are phenotype Ag(a+) and those homozygous for *Ag* are phenotype Ag(a−).

Since the discovery of the first precipitin-containing serum, nearly 20 others have been found. (Blumberg, 1963; Blumberg and Riddell, 1963; Blumberg, Alter, Riddell, and Erlandson, 1963). By immunological and epidemiological studies, it has been shown that these can distinguish several different specificities. The genetics of two of these specificities have been studied in some detail, but the exact genetic interrelationship of the two has not been determined. An interesting problem in the study of the genetics became apparent as more of the antisera were discovered. For several of them (*i.e.*, New York and J.B.) nearly all the sera from U.S. whites and Negroes were reactors. The probable explanation for this is that patients who lack an antigen common in the donor population are likely to develop an antibody against the common antigen as a result of repeated transfusions. As a consequence of this, it is necessary to conduct the family and twin studies in populations where the antigens are less common so that suitable families for testing a genetic hypothesis can be more readily found. The results of our surveys in several populations are shown in table 3. These indicate that American Indian and Oceanic populations are most suitable for genetic studies with several of the antisera.

Having determined that the antigenic difference was inherited, the next step was to determine what, if any, was the association with disease.

For the first antibody (anti-Ag(a+)), there was no indication of any direct association of any of the phenotypes with overt disease. There is less information about the other antisera, but further investigations are in progress.

A priori it was difficult to know which diseases to study, but it was considered logical to investigate illnesses in which the β-lipoprotein

TABLE 3.—*Reactors with antisera in different populations*

	C. de B. Anti—A		New York Anti—B		J. B.		A. DiB		S. L.	
	Number	Percent+	Number	Percent+	Number	Percent+	Number	Percent+	Number	Percent+
Whites (U.S.A.)	120	59	164	100	168	98	144	77	157	97
Micronesians (Rongelap)	194	98	185	45	54	98	45	24	50	24
Naskapi-Montagnais (Labrador)	234	97	103	85	234	93	91	73	93	95

was known to be abnormal. These include coronory artery diseases, arteriosclerosis, diabetes, a variety of familial lipodystrophies, and others. The results of several preliminary investigations will be given here.

1. *Coronary artery disease.*

Through the kindness of Dr. T. R. Dawber, Heart Disease Epidemiology Study, Framingham, Mass., we were able to obtain serum specimens suitable for a preliminary investigation. The design of the Framingham study has recently been discussed by Dawber *et al.*, (1963). Briefly, approximately 13 years ago, a health survey of a large portion of the adult population of Framingham was conducted. Most of these people have been followed and those who have developed coronary artery disease, as well as other forms of heart disease, have been detected. Dr. Dawber sent to us sera from 69 individuals who had developed coronary artery disease and sera from 67 individuals of the same sex and approximately the same age who had not. Most of these specimens were drawn several years before the testing, but the test and control sera have approximately the same history of storage and handling. All the sera were phenotyped using three of our antisera.

The results are shown in table 4. There is no significant difference between the control group and the patients for any of the antisera which were studied. The frequencies of reaction with New York antiserum and J.B. antiserum are lower than those previously reported, but this is probably due to the long storage of the sera (Blumberg and Riddell, 1963). However, since both the patient and the control samples have been handled in the same manner, it is unlikely that this would have significantly altered the results. Although this study could hardly be said to rule out the possibility of an association, it does suggest that there might be more profitable avenues to follow.

2. *Diabetes.*

For the preliminary study, American Negro patients with known diabetes were casually selected from the Diabetes Clinic at the Washington Hospital Center over a 4-month period. For controls, sera were obtained from Negro employees of the Center during the

TABLE 4.—*Association of Ag systems with disease*

	Human anti-human antisera								
	C. de B.			New York			J.B.		
	Anti−Ag(a+)			Anti−Ag(b+)					
	Total	Percent POS	?	Total	Percent POS	?	Total	Percent POS	?
Coronary artery disease	68	60.3	1	67	88.1	3	57	75.4	13
Control	66	59.1	1	66	77.3	2	56	76.8	11
Diabetes	208	71.6	0	205	96.6	2	205	97.1	2
Control	166	55.4	1	159	98.1	2	154	94.8	5
Rheumatic fever	165	83.6	0	37	100	0	35	100	0
Control	152	82.2	0	34	100	0	34	100	0

routine preemployment medical examination over the same period of time. There was a higher frequency of C. de B. reactor (Ag(a+)) individuals among the diabetics than among the controls, and this difference is significant ($\chi^2=10.4$, $P<.001$). The differences using the other antisera were not significant (table 4). We are now evaluating the meaning of this finding and have received some comfort in our agony from the words of the late Prof. Sir Ronald A. Fisher. Following a lecture on correlations at the National Institutes of Health several years ago, I excitedly presented to him some data which showed a definite correlation between certain inherited traits. When I asked him if he could comment on the value of these results, he said "Blumberg, you have something they can never take away from you."

Before making any conjectures on the meaning of these results, it is clear that additional studies should be done. If these show the same striking difference, then a further examination of the genetic and biochemical significance of the correlation would be warranted.

3. *Rheumatic fever.*

Susceptibility to rheumatic fever is not generally associated with abnormalities of lipoproteins. However, there have been some studies which indicate that the serum lipid-containing fractions are abnormal in this disease. Our major reason for studying this disease was the ready availability of sera from patients with rheumatic fever and their spouse controls. These were presented to us through the kindness of Dr. Thomas D. Dublin and his coworkers. The method of ascertainment and selection of patients and controls is described elsewhere in this conference. There is no difference in the frequency of reactors with C. de B. in the patient and control group. To conserve antiserum, a smaller number of patients and controls were studied with the other two antisera, and they also showed no differences (table 4).

4. *Lipodystrophies.*

The sera of patients with a variety of lipodystrophies were presented to us through the courtesy of Dr. D. Fredrickson. In one of these, the C. de B. antiserum reacted with the patient's serum to form a precipitin line; the line did not stain with Sudan Black, but did do so with Azocarmine. This suggests that the antigen is not a lipoprotein. We have seen only one other instance in the thousands of sera tested with C. de B. in which the antigen did not contain a lipid-staining material (Blumberg, Alter and Riddell, 1963), and in this other case it was not possible to examine the subject. We are continuing our investigation to determine if this rare phenotype is in any way connected with the patient's illness.

Patients with acanthocythosis lack all or nearly all serum beta-lipoprotein. We have studied four such patients and the immediate families of two of them. There is no convincing evidence at present that a gene for absent beta-lipoprotein exists. However, in two of the cases we found that the serum of these patients formed a precipitin line with a lipid-containing protein present in the serum of some but not all normal persons. One of these abetalipoproteinuria patients had received a single transfusion ten years prior to the examination, and the second had never had a transfusion (Blumberg, 1963). This suggests that if in fact this material is an antibody, the patients developed it as a result of immunization with some nonhuman material. The other possibility is that some form of "autoimmunity" is involved.

SUMMARY

A method of genetic analysis of disease susceptibility using polymorphic traits is described. The beta-lipoprotein Ag system is used to illustrate the technique.

REFERENCES

1. Allison, A. C., and Blumberg, B. S. 1961. An isoprecipitation reaction distinguishing human serum protein types. *Lancet* 1: 634–637.
2. Aschaffenburg, R., and Drewry, J. 1955. Occurrence of different beta-lactoglobulins in cow's milk. *Nature* 176: 218–219.
3. Beutler, E. 1959. The hemolytic effect of primaquine and related compounds. A review. *Blood* 14: 103–139.
4. Blumberg, B. S. 1963. Iso-antibodies in humans against inherited serum low-density beta-lipoproteins, the Ag system. *Ann. N.Y. Acad. of Sciences* 103: 1052–1057.
5. Blumberg, B. S. 1961. Inherited susceptibility to disease. Its relation to environment. *Arch. of Environ. Health* 3: 612–636.
6. Blumberg, B. S., Alter, H. J., Riddell, N. M., and Erlandson, M. 1964. Multiple antigenic specificities of serum lipoproteins detected with sera of transfused patients. *Vox Sanguinis* 9: 128–145.
7. Blumberg, B. S., Bernanke, A. D., and Allison, A. C. 1962. A human lipoprotein polymorphism. *J. Clin. Investigation* 41: 1936–1944.

8. Blumberg, B. S., Dray, S., and Robinson, J. C. 1962. Antigen polymorphism of a low density β-lipoprotein. Allotypy in human serum. *Nature* 194: 656–658.
9. Blumberg, B. S., Kuvin, S. F., Robinson, J. C., Teitelbaum, J. M., and Contacos, P. G. 1963. Alterations in haptoglobin levels. *J. Amer. Med. Assoc.* 184: 1021–1023.
10. Blumberg, B. S., and Riddell, N. M. 1963. Inherited antigenic differences in human serum beta-lipoproteins. A second antiserum. *J. Clin. Invest.* 42: 867–875.
11. Blumberg, B. S., and Robbins, J. 1960. Thyroxine-serum protein complexes: single dimension gel and paper electrophoresis studies. *Endocrin.* 67: 368.
12. Blumberg, B. S., and Tombs, M. P. 1958. Possible polymorphism of bovine α-lactalbumin. *Nature* 181: 683–684.
13. Dawber, T. R., Kannel, W. B., and Lyell, L. P. 1963. An approach to longitudinal studies in a community. *Ann. N.Y. Acad. of Sciences* 107: 539–556.
14. Ford, E. B. 1956. Genetics for Medical Students. London, Methuen.
15. Gartler, S. M. 1959. A metabolic investigation of urinary β-aminoisobutyric acid excretion in man. *Arch. Biochem. and Biophys.* 80: 400–409.
16. Gartler, S. M., and Gandini, E., Ceppellini, R. 1962. Glucose-6-phosphate dehydrogenase deficient mutant in human cell culture. *Nature* 193: 602–603.
17. Giblett, E. R. 1958. Js, a "new" blood group antigen found in Negroes. *Nature* 181: 1221–1222.
18. Grabar, P., and Williams, C. A., Jr. 1955. Méthode immuno-electrophorétique d'analyse de mélanges des substances antigéniques. *Biochem. et Biophys. Acta* 17: 67–74.
19. Hirschfeld, J., and Beckman, L. 1960. A new group-specific serum system (GC groups) in relation to blood and serum groups. *Acta Genet.* 10: 48–53.
20. Hockwald, R. S., Arnold, J., Clayman, C. B., and Alving, A. S. 1952. Status of primaquine. 4. Toxicity of primaquine in Negroes. *J. Amer. Med. Assoc.* 149: 1568–1570.
21. Jayle, M. F., Boussier, G., and Badin, J. 1952. Electrophoresis de l'haptoglobine et de son complex hemoglobinique. *Bull. Soc. Chim. Biol.* 34: 1063–1069.
22. Kalow, W. 1959. Cholinesterase types. In CIBA Found. Symposium. Biochemistry of Human Genetics. Boston, Little, Brown and Co.,
23. Levine, P., Robinson, E. A., Layrisse, M., Arends, T., and Dominques Sisco, R. 1956. The Diego blood factor. *Nature* 177: 40–41.
24. McDonough, J. R., Hames, C. G., Greenberg, B. G., Griffin, L. H., Jr., and Edwards, A. J., Jr. 1963. Observations on serum cholesterol levels in the twin population of Evans County, Georgia. *Circulation.* In press.
25. Parker, W. C., and Bearn, A. G. 1963. Control gene mutations in the human haptoglobin system. *Nature* 198: 107–108.
26. Robinson, J. C., Blumberg, B. S., Pierce, J. E., Cooper, A. J., and Hames, C. G. 1963. Studies on the inherited variants of blood proteins. II. Familial segregation of transferrin $B_{1-2}B_2$. *J. Lab & Clin. Med.* 62: 762–765.
27. Smith, C. A. B. 1956. A test for segregation ratios in family data. *Ann. Hum. Genet.* 20: 257.
28. Smithies, O. 1959. An improved procedure for starch-gel electrophoresis. Further variations in the serum proteins of normal individuals. *Biochem. J.* 71: 585–587.

29. Smithies, O. 1955. Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults. *Biochem. J.* 61: 629.
30. Smithies, O., and Connell, G. E. 1959. Biochemical aspects of the inherited variations in human serum haptoglobins and transferrins. In CIBA Found Symposium: Biochemistry of Human Genetics. Boston, Little, Brown and Co.
31. Smithies, O., Connell, G. E., and Dixon, G. H. 1962. Chromosomal rearrangements and the evolution of haptoglobin genes. *Nature* 196: 232–236.
32. Smithies, O., and Hiller, O. 1959. The genetic control of transferrins in humans. *Biochem. J.* 72: 121.
33. Steinberg, A. G., and Giles, B. D. 1959. A genetically determined human serum factor selected by its effect on a mating reaction in yeast. *Amer. J. of Human Genetics* 11: 380–384.

PART IV—Continued

Regional Variations in Morbidity and Mortality

Regional Variations in the Incidence of Spina Bifida

By David Hewitt, M.A., *Department of Social Medicine, Oxford University, Oxford, England.*

The contributors to this symposium, and especially to the workshop, were enjoined "to emphasize the experimental design of (their studies) with . . . particular emphasis on how (they) proposed to define familial factors and distinguish between the genetic and the non-genetic. . . ." For the purpose of such a methodological discussion the study which I have to describe is a nonstarter. It is merely a piece of descriptive epidemiology, based entirely on official population statistics, and thus employing no information—genetic *or* nongenetic—about the individual families in which cases of spina bifida occur. However, it did give rise to a somewhat cloudy hypothesis concerning a genetic aetiology for spina bifida, and possibly for other related malformations of the central nervous system. It will be for others to judge whether this hypothesis deserves to be taken seriously and, if so, to suggest an appropriate design for a further study to test the hypothesis.

So far from starting with a problem and then setting up appropriate machinery for studying it, I started with certain machinery—essentially a desk calculator and a medical library—and then looked for a problem to which this equipment might usefully be applied. A great abundance of second-hand epidemiological data is available in the publications of the National Office of Vital Statistics and in particular in their series of annual reports on the "Vital Statistics of the United States" (National Office of Vital Statistics, 1947–62). This suggested the feasibility of a study of regional—specifically of interstate—variations in mortality. To avoid the problems of interpretation which arise when there is a possibility of selective migration of sick persons, it seemed prudent to focus on mortality occurring very early in life. Two other desiderata affecting the choice of a particular topic were:

(i) that the statistics should indicate more than trivial differences between States, and

(ii) that the reported differences should not be artefacts produced by interstate differences in diagnostic practice.

Tables 1 and 2 show that, according to certain criteria, spina bifida meets both these requirements.

In table 1 are given values of a statistic intended to measure the amount of interstate variation in the mortality of white infants during the decade 1950–59. This statistic is analagous to the coefficient of variation in common use, but has been defined in such a way [1] as to measure only the relatively permanent differences between States, excluding differences due to chance or other short-term factors. It will be seen that, relatively speaking, the infant death-rates from spina bifida are more than twice as variable from State to State as are the infant death-rates from all causes, and more than five times as variable as the infant mortality from congenital malformations as a whole.

TABLE 1.—*"Reduced" coefficient of variation of mortality rates (based on white infants of each State and of the District of Columbia, 1950–54 and 1955–59)*

	Percent
All causes:	
Ages 0 months to 11 months	12.5
Ages under 1 month	8.8
Ages 1 to 11 months	27.9
Congenital Malformation:	
All types	6.0
Spina bifida	31.9
Circulatory system	5.8
Other malformations	6.7

The credibility of this reported interstate variation of spina bifida mortality was tested by correlating the State death-rates with an index of the quality of death certification in each State. The construction of this index has been described elsewhere (Hewitt, 1963). As would be expected of such an index it tends (see table 2) to be low where post-neonatal infant mortality as a whole is high. It also has a significant positive correlation with the reported mortality from certain types of malformation, suggesting that, in some States at least, the reported rates may be low merely because of inadequate diagnosis or certification. However, for spina bifida the correlation is as near zero as one could wish.

TABLE 2.—*Correlation between an index of diagnostic quality and the reported mortality of white infants in each State, 1950–59*

All causes:	
Ages 1 to 11 months	−0.481
Congenital malformation:	
Spina bifida	−0.008
Circulatory system	+0.081
Hydrocephalus*	+0.343
Other malformations*	+0.428

*Including some deaths after the age of 1 year.

[1] $\frac{\text{(Co-variance of 1950-54 rate with 1955-59 rate)}^{\frac{1}{2}}}{\text{Mean of 1950-54 and 1955-59 rates}}$

Hence spina bifida is a "good" subject for armchair epidemiology, because the geographical variation is both large and credible.

The next question is, how can this geographical variation be summarized? Following the precedent set by Wesley's (1960) study of congenital malformations it was decided in the first instance to correlate the State mortality rates with measures defining the geographical position of the States, thus testing the possibility that the risk of spina bifida might vary in continuity over some or all of the country. In Wesley's work a prior interest in the possible relevance of cosmic-ray energy flux caused him to use, as his one and only measure of position, the magnetic latitude; that is, a particular joint function of geographical latitude and geographical longitude. In the present case it seemed preferable to give independent consideration to latitude and to longitude. The result was striking. Spina bifida mortality in the United States has *no* correlation with latitude ($r = -0.046$). On the other hand, longitude of the State capital is a very good predictor of spina bifida mortality ($r = -0.839$), and may be said to "account for" just over 70 percent of the variance between States, or more than twice as much as remains to be accounted for by all those factors, including sampling error, which are not correlated with longitude. The regression equation estimates that between the extremes of longitude spanned by the conterminous United States, risk of fatal spina bifida differs by a factor of nearly three.

Figure 1 shows how this relationship appears on the map. This map differs from the one drawn by Alter (1962) by incorporating 4 more years' data and, more importantly, by excluding nonwhites, whose low average rates and disproportionate concentration in the southern States complicate the picture. When only white infants are considered, this picture is rather simple. Of the 24 States with rates below the national median, 20 lie entirely west of the Mississippi, and 3 of the 4 exceptions are contiguous with the low area. Only Connecticut is "out of place".

What happens outside this range of longitude? It would be a naive extrapolation which suggested that the rate in Alaska should be lower than for any of its 48 seniors and that in Hawaii the rate should be lower again, but both these suggestions turn out to be correct! Moreover, Puerto Rico, which lies somewhat east of Maine, has a rate higher than all but one of the present 50 States.

The evidence for east-west variation is greatly strengthened by data from Canada (Dominion Bureau of Statistics, 1957–61). This is shown in table 3, where the Provinces have been arranged in order from east to west. Except that the rate for Quebec is too high, this arrangement also corresponds to their ranking by spina bifida mortality.

The cases of spina bifida which enter into the numerators of these death-rates are far from being all that occur. Some cases will die

735-712 O—65——20

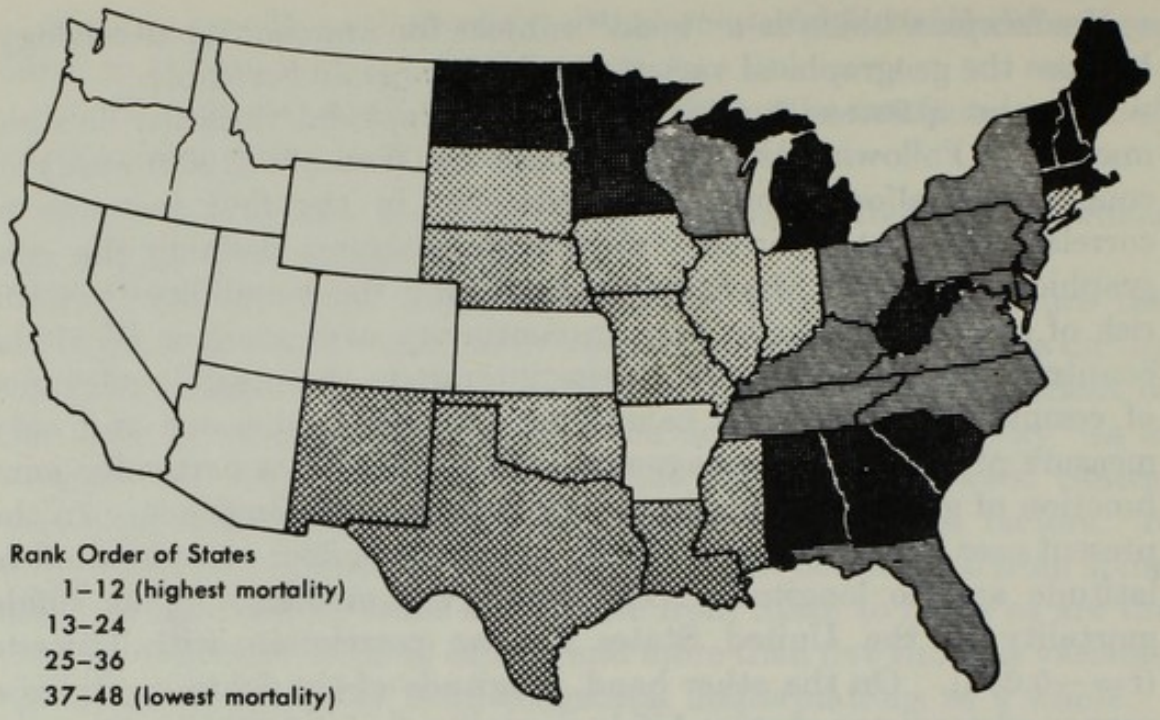


FIGURE.—Mortality of white infants attributed to spina bifida and meningocele in each State of the U.S.A., period 1950–59.

TABLE 3.—*East-west variation in Canada of mortality attributed to congenital malformation*

(Per 10,000 live births, 1955–59)

	Spina Bifida	Other Malformations
Newfoundland and Maritimes	9.8	51.1
Quebec	11.7	57.4
Ontario	8.0	50.0
Manitoba	6.9	45.8
Saskatchewan	5.3	45.1
Alberta	4.6	48.0
British Columbia	4.1	44.6

in utero and not appear in the cause-of-death statistics, while others which do appear in these statistics will be placed in a different category because of coexistent hydrocephalus. It has to be allowed that inclusion of these lost cases might modify the geographical pattern to some extent. But there is reason to think that they would reinforce rather than weaken the general east-west trend. Thus, post-natal mortality certified as due to hydrocephalus is also heavier in eastern than in western States and—a significant detail—this is especially true of females. Also the data on medical causes of stillbirth in the Provinces of Canada (Dominion Bureau of Statistics, 1957) indicate that intra-uterine death attributable to spina bifida is commoner in Quebec than in Ontario, and commoner in Ontario than in any of the four western Provinces.

No parallel can be traced in the mortality of nonwhites. Analysis of the data on nonwhites is complicated by the fact that American Indians, Negroes and other nonwhites have reported mortalities from

spina bifida in approximately the ratio 6:3:1. These contrasting groups have very different geographical distribution, but cannot always be separated out in the statistics. A striking difference between the white and Negro infants dying of spina bifida is that among the former (in Europe as well as in America) there is a great preponderance of girls; among Negroes there is a more normal and distinctly masculine sex-ratio. When such a basic epidemiological characteristic as the sex-ratio among affected individuals differs so markedly between the races, one should perhaps not expect the geographical patterns to be the same.

At first look the east-west gradient of the white death-rates may seem nonsensical. The strong correlation with *latitude* found in multiple sclerosis strikes one as quite reasonable, even while the reason for it remains unknown, because of the variety of conceivably relevant physical factors which also vary with latitude (Acheson, Bachrach and Wright, 1960). But no east-west polarity exists in nature. Hence, if we are to make any sense of the finding, we need to turn up some factors in the history or culture of the human population which—in this continent[2]—could express themselves as differences between east and west. Lacking any data of a more appropriate type, we can only pursue the search for such factors by correlating the State death-rates with whatever other average characteristics of the State populations happen to be known. In this way we may hope to find a principle of classification of the States which will separate out those with high and low mortality rates from spina bifida, which will necessarily correspond to some extent to the classification by longitude, but which will be more directly meaningful than is degrees west. Table 4 summarizes some of the classifications which have been tried so far. In each column of the table there appear 4 figures; each figure is the average of 12 State mortality rates from spina bifida (Alaska, Hawaii and the District of Columbia have been excluded), and each group of 12 States has been composed according to the ranking of the States by the characteristic named, with the highest on line A and the lowest on line D.

In the first column the States have simply been ranked by spina bifida mortality itself, to provide a standard of comparison for the other columns of the table. Column 2 recapitulates the finding on longitude.

One fairly safe generalization about western areas of the United States is that they have been under white settlement for a relatively short time. In case this should be thought relevant, column 3 shows the effect of classifying the States by a rough index of "community age", namely, the order of admission of the States to the Union.

---

[2] In Europe, too, some of the variations in frequency of spina bifida as well as of anencephalus conform to an east-west pattern, but with the higher levels in the west (Penrose, 1957; Lamy and Frézal, 1959).

TABLE 4.—*Mortality rates of white infants from spina bifida averaged over the States lying in the first, second, third and fourth quartiles on certain rankings*

	1	2	3	4	5	6	7
A	6.05	5.50	4.94	4.47	4.34	5.44	5.44
B	4.58	4.89	5.00	5.39	3.96	4.63	3.91
C	3.74	3.78	3.78	4.17	4.10	3.82	3.89
D	2.46	2.65	3.11	4.42	4.42	2.94	3.60

Key to rankings:

1. By recorded mortality from spina bifida.
2. By longitude of State capital.
3. By order of admission to the Union.
4. By ratio of 1890 to 1960 population.
5. By percentage of 1959 mothers foreign-born.
6. By percentage of naval recruits lifetime residents of State.
7. By a feature of the birthweight distribution (see text).

Pursuing the same line of thought, the newer communities tend to be those whose populations have expanded most rapidly in the last two generations. Accordingly, column 4 uses a classification by the ratio between the 1890 and the 1960 population. This is fairly successful in picking out the States with the lowest rates.

One element in the 20th-century population expansion has been new immigration from overseas, but, among the parents of today's children, first-generation immigrants are not particularly numerous, and they do *not* make up any larger proportion in the western States generally, or in the States with low spina bifida mortality (see column 5).

Probably a more important factor in the differential growth-rates of State populations has been the westward tide of internal migration, in consequence of which the young adult populations of western States contain a relatively high proportion of people who have moved some distance from their parental homes. Adapting figures given by Edwards and Palmer (1963) for the proportion of white naval recruits who were lifetime residents of the State in which they enlisted, we obtain the ranking of States shown in column 6. This index of the "rootedness," as against mobility, turns out to be almost as good a predictor of spina bifida mortality as is the longitude. Why should such an index have any predictive value? One possible explanation might be that matings in the more mobile populations more frequently involve outcrossing between stocks with contrasting allelic frequencies at a number of loci, so that the offspring tend to have a higher proportion of loci in a heterozygous state. According to one of the theses put forward by Lerner (1954), and previously discussed in relation to human malformations by Neel (1958), this could reduce the risk of abnormal morphological development. One attraction of Lerner's thesis in this context is that it suggests a perspective on regional and temporal variations of spina bifida much broader than that of America in the 1950's. Though the rates on the Atlantic seaboard may have been high by American standards, yet they were lower, for example, than contemporary rates in any major section of the British Isles, from which much of the parental stock (in particular, that of Canada's

Maritime Provinces) was drawn. We must also find a place in the picture for the downward secular trend of spina bifida mortality (MacMahon *et al.*, 1951; Gittlesohn and Milham, 1962; Hewitt, 1963). It may be that what we have to explain is not an excess of affected individuals in certain parts of North America, but a falling incidence in the white population of the world, with particularly low rates in those regions where there is the most intermarriage between formerly separate subpopulations such as the European national groups.

What independent evidence is there of greater outbreeding in the western States and Provinces than elsewhere, or of any connection between this and frequency of spina bifida? I shall offer two findings consistent with these ideas. The first was suggested by a paper of Morton's (1958) in which he stated that on theoretical grounds relatively intense inbreeding would be expected to reduce the mean and increase the variance of the distribution of birthweights. As it happens, this distribution is known for liveborn white infants of each State through the years 1955–59. Because scatter towards the lower end of the birthweight distribution could be due to a variety of irrelevant factors, I decided to try a measure based on the scatter at the upper end of the distribution, namely the ratio

$$\frac{\text{number of live births weighing 5,001 g or more}}{\text{number of live births weighing 4,501 g or more}}.$$

This somewhat esoteric index appeared to perform rather well (see column 7 of table 4), in that it grouped the States into 4 sets whose average mortality rates were in exactly the predicted rank order ($P=1/24$).

The second and more direct piece of evidence relates to Canada, and is derived from a special tabulation of legitimate live births in 1950 according to the "national origin" of the parents in combination (Dominion Bureau of Statistics, 1955, table 34). Matings between persons of the same "national origin" ranged from a low of 37.5 percent of all matings in British Columbia up to a high of 91.7 percent in Quebec. As previously shown, these two Provinces were respectively the lowest and the highest for spina bifida mortality. Between these extremes the five other divisions of Canada shown in table 3 preserve almost the same ranking for spina bifida as for this measure of endogamy, the only discrepancy being the reversal of Manitoba and Ontario.

The composite hypothesis suggested on the basis of this study is therefore as follows:

(i) that the incidence of spina bifida in white populations is influenced by the average degree of homozygosity; and

(ii) that both the downward time-trend of recent years and the downward geographical trend as one proceeds westward across America are due to greater outbreeding.

This is clearly akin to the hypothesis of recessive inheritance proposed by Polman (1950) on the basis of certain inbred Dutch communities in which he considered there was an excessive frequency of several types of central nervous system malformation.

One followup to the present study which may seem worthwhile, whether or not the suggested hypothesis is taken seriously, is to determine whether anencephalus is also commoner in the east than in the west of the United States. A positive or a negative finding would be equally interesting. A second type of followup would be to enumerate the number of "national origins" represented in the immediate ancestry of malformed and control infants, possibly along the lines of the study already proposed by Bresler (1962) for testing a hypothesis about the total fertility of marriages between members of different religious groups.

REFERENCES

1. Acheson, E. D., Bachrach, C. A. and Wright, F. M. (1960). Some comments on the relationship of the distribution of multiple sclerosis to latitude, solar radiation, and other variables. Acta psychiat. scand., Suppl. 147, *35*, 132–147.
2. Alter, M. (1962). Anencephalus, hydrocephalus and spina bifida. Arch. Neurol. (Chic.), *7*, 411–422.
3. Bresler, J. P. (1962). The relationship between the fertility patterns of the F_1 generation and the number of countries of birth represented in the P_2 generation. Amer. J. Phys. Anthropol., *20*, 509–514.
4. Dominion Bureau of Statistics (1955). Vital Statistics for the Year 1950. Queen's Printer and Controller of Stationery, Ottawa.
5. Dominion Bureau of Statistics (1957). Causes of Stillbirth 1949–55. Reference Paper No. 79. Queen's Printer and Controller of Stationery, Ottawa.
6. Dominion Bureau of Statistics (1957–61). Vital Statistics for the Years 1955–59. Queen's Printer and Controller of Stationery, Ottawa.
7. Edwards, P. Q. and Palmer, C. E. (1963). Nationwide histoplasmin sensitivity and histoplasmal infection. Publ. Hlth. Rep. (Wash.) *78*, 241–259.
8. Gittlesohn, A. M. and Milham, S. (1962). Declining incidence of central nervous system anomalies in New York State. Brit. J. prev. soc. Med., *16*, 153–158.
9. Hewitt, D. (1963). Geographical variations in the mortality attributed to spina bifida and other congenital malformations. Brit. J. prev. soc. Med., *17*, 13–22.
10. Lamy, M. and Frézal, J. (1959). *In* Coll. int. malformations congénitales de l'encéphale. Paris: Masson.
11. Lerner, I. M. (1954). Genetic Homeostasis. Wiley, New York.
12. MacMahon, B., Record, R. G. and McKeown, T. (1951). Secular changes in the incidence of malformations of the central nervous system. Brit. J. soc. Med., *5*, 254–258.
13. Morton, N. E. (1958). Empirical risks in consanguineous marriages: birth weight, gestation time and measurements of infants. Amer. J. Hum. Genet., *10*, 344–349.

14. National Office of Vital Statistics (1947–62). Vital Statistics of the United States for the Years 1945–49. U.S. Government Printing Office, Washington.
15. Neel, J. V. (1958). A study of major congenital defects in Japanese infants. Am. J. Human Genet., *10*, 398–445.
16. Penrose, L. S. (1957). Genetics of anencephaly. J. Ment. Defic. Res., *1*, 4–15.
17. Polman, A. (1950). Anencephaly, spina bifida and hydrocephaly. Genetica, *25*, 29–78.
18. Wesley, J. P. (1960). Background radiation as the cause of fatal congenital malformation. Int. J. Radiat. Biol., *2*, 97–116.

Vital Record Incidence of Congenital Malformations in New York State

By A. M. Gittelsohn, Ph. D., M.P.H., and S. Milham, Jr., M.D., M.P.H., *Director of Health Statistics and Development Consultant, New York State Department of Health,* Albany, N.Y.*

A concomitant of the growth of the fields of human genetics, experimental teratology and cytogenetics, has been an increased interest in the incidence of congenital malformations in human population groups. As stated causes of death *per se*, congenital defects account for almost one out of five deaths in infancy and early childhood. Malformations are also assigned an important role in abortion and fetal loss. The recent thalidomide experience suggests that the interval of more than one year between appearance of drug-associated cases and removal of the product from distribution might have been reduced had incidence data been available on a central basis. The wide geographic and temporal distribution of phocomelia cases among the total population of births reduced the opportunity for any one physician to observe a significant number of cases.

Since 1940, the New York State birth certificate has contained an entry for recording congenital malformations. Analysis of the data has been sporadic and no routine procedure was developed for tabulating the data. The present investigation was undertaken from several viewpoints. The objectives were to delineate the usefulness and limitations of information on congenital defects obtained through the vital record system; for those defects reported with sufficient quality, to describe patterns in incidence relative to available geographic and demographic characteristics; and to record the "background" incidence against which secular changes could be measured.

The population of events consists of over 2 million live births and fetal deaths registered in upstate New York or occurring to residents

*Present address of senior author: Department of Biostatistics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Md.

of the jurisdiction during the 11 years between 1950 and 1960. (Upstate refers to New York State exclusive of New York City.) Through a routine procedure, birth records are updated by death information on infants and children up to 5 years of age. The combined birth and death data are available for producing survivorship tables and mortality rates in the perinatal, infancy, and the early childhood periods of life. Malformation cases hence can be ascertained through birth and/or death reports. Because of the inadequacy of the WHO Statistical List for coding malformations, all fetal death certificates reporting a congenital defect were searched and recoded according to a more specific classification. Live birth and death records with malformations of the CNS or with multiple malformations were similarly treated.

Completeness of case ascertainment is a central problem in epidemiologic studies of incidence rates. Rate in this connection is defined in terms of the relative frequency of specified events within a defined population during a given time period. Registration is required for all products of conception of over 20 weeks gestation. Thus, early fetal deaths and abortions are not available for observation and analysis. The completeness of birth registration is estimated to be well above 99 percent.

By contrast, the recording of malformation data on birth certificates is highly variable, depending primarily on recognizability of the defect by external examination. In order to assess completeness, a number of tests were carried out. Beginning with a group of children receiving treatment for specific congenital defects under the State Aid Program, examination of corresponding birth records revealed a wide range of reporting variability (table 1). As anticipated, internal defects were generally not noted at the time of birth while external defects tended to be picked up by the system. Eighty percent of 240 spina bifida cases, as contrasted with less than 10 percent of 226 circulatory system malformations, were recorded. The incompleteness would tend to be overstated under such circumstances as the children with the more severe forms of the defect, incompatible with extended postnatal survival, would not enter the program.

TABLE 1.—*Children Treated in State Aid Program by Type of Defect and Vital Record Statement*

Type of defect	Number in State Aid program	Defect noted on birth record	Noted death record only	Percent recorded
Spina bifida	240	188	4	79.9
Cleft lip and/or palate	1,096	817	3	74.8
Hypospadias	97	43	----	44.3
Club foot	68	30	----	44.1
Polydactyly	52	21	----	40.4
Hydrocephalus	79	12	15	34.3
Syndactyly	19	6	----	31.6
All circulatory defects	226	----	19	8.4

As one of the more important malformations of the central nervous system, anencephalus has long been recognized; its appearance is so marked that it is unlikely not to be noted. A certain number of cases undoubtedly are not recorded on the birth record, and others are described under generic terms such as "monster" and are thereby lost. In the course of developing sibship data for probands with anencephalus or spina bifida the diagnostic indexes of the three maternity hospitals in one city were exhaustively searched (Milham, 1962). In tracing the birth records of the 150 cases so ascertained, it was found that more than 96 percent had the defects noted. In so far as the diagnostic indexes are complete, the birth certificates for at least one community accurately reflect their content for the two neural tube anomalies.

In a similar approach to evaluating reporting of clefts of the lip and palate, the diagnostic indexes of three large teaching hospitals in the State were searched (Milham, 1963). All cases reported among the 79,536 births occurring in the three institutions between 1950 and 1960 were collected. Comparison of the two information sources revealed the following:

Cases of Cleft Lip and/or Palate by Type of Ascertainment

	Number of cases	Cases per 100,000 births
Birth certificate only	104	131
Hospital indexes only	117	147
Both sources	143	180
Total births	79,536	

On the basis of both sources, the incidence was 180 cases per 100,000 births. This is a minimum figure, since there may be cases not recorded on both birth certificate and diagnostic index. Ascertainment by birth certificate alone would have led to a 27 percent lowered incidence, while ascertainment by hospital indexes alone would have resulted in a loss of 18 percent of the cases. Since the three institutions were teaching hopitals with active obstetric, pediatric, and plastic surgery services, the record systems would be expected to be above average and the agreement between birth certificates and hospital information probably represents an overstatement of the statewide pattern.

In addition to the problem of completeness of ascertainment of cases through vital registration is the question of specificity of the morphologic description of the malformation. For clefts of the lip and palate this is a problem of particular importance. In carefully constructed case series, it has often been noted that isolated cleft palate behaves differently from cleft lip with or without palatal involvement. Sex incidence for the former defect is predominantly female, while for the

latter it is predominantly male. Experience has shown that the statement "cleft palate" on the birth record may or may not exclude lip involvement. A similar situation pertains to the statement "hare lip." Anencephalus may be variously described under such terms as acephalus, acrania, aplasia of brain, or monster. Birth records are not designed for obtaining precise anatomic detail on which specific classifications can be traced.

A further complication is introduced by the use of the present WHO Statistical Classification. The code category No. 750 includes such diverse entities as anencephalus, hemicephalus, exencephalus, syncephalus and monster, not otherwise stated. Adactylism is lumped with polydactylism, amelia and phocomelia. To this end, a new classification system allowing for greater specificity had to be developed and each record recoded thereby.

In general, the vital record system is a useful mechanism only for ascertainment of gross external malformations readily apparent at birth. A tentative listing of those defects which we feel are recorded with sufficient frequency and precision to be useful for certain limited purposes, is as follows:

Anencephalus	Conjoined twins
Spina bifida	Exomphalos
Imperforate anus	Polydactyly
Exstrophy of bladder	Gross deformities of limbs

In general, the quality of information on vital records is highly variable. The major virtue of the system is to be found in the completeness of birth and/or death registration from a population standpoint. Beyond the simple fact of birth or death and statements of identification, there is a lack of uniformity in the underlying observations pertaining to pregnancy, labor, and the condition of the child at birth.

VITAL RECORD INCIDENCE OF CONGENITAL MALFORMATIONS

The possibilities for analyses of incidence data derived from vital records include spatial and temporal distributions, survivorship, and case rates by parental ages, parity, sex, and race. The present purpose is to summarize briefly a few of the findings.

During the period 1950–60, in upstate New York, a total of 26,728 children with one or more congenital malformations were recorded on fetal death, live birth and/or death certificates. Counting children rather than malformations, the total incidence was 1,305 cases per 100,000 births. The defects ranged in severity from anencephaly to polydactyly. When two or more malformations were coexistent

in a single individual, coding assignment was to the more severe. Information on multiple malformations was preserved for future analysis.

Table 2 shows the number of cases by source of vital record information. For the CNS defects, the majority of cases was reported on birth records. Over one-half of the anencephalus and hydrocephalus cases were stillborn and of those born alive, most died during the first year. For spina bifida, a smaller proportion was stillborn (13 percent), and over one out of four were not known to be dead after 5 years. By contrast, information on internal malformations involving the circulatory, digestive, G.U., or respiratory system was largely ascertained through the death certificate only, the corresponding birth record containing no indication of the defect.

TABLE 2.—*Number of Malformations by Type and Source of Data, Upstate New York, 1950–60*

	Total cases	Fetal deaths	Deaths first year	Deaths 1–4 years	Not known dead	Percent fetal deaths	Percent on D.C. only
All malformations	26,728	2,853	8,914	736	14,225	10.8	16.2
CNS:							
Anencephalus	1,504	912	585	2	5	60.5	3.9
Spina bifida	2,025	272	1,103	92	558	13.4	5.1
Hydrocephalus	1,580	853	526	95	106	54.0	16.8
Other CNS	711	220	280	56	55	30.9	21.0
Circulatory:							
Tetralogy of Fallot	95	0	73	19	3	--------	77.8
Patent ductus arter	190	7	170	7	6	3.6	73.6
IV septal defect	266	24	216	22	4	.9	81.2
IA septal defect	200	14	167	11	8	7.0	68.0
Other	3,112	90	2,553	145	324	2.9	66.1
Digestive:							
Cleft lip/palate	2,259	24	261	25	1,947	1.1	.1
Pyloric stenosis	71	--------	66	1	4	--------	91.5
Imperforate anus	300	5	104	8	183	1.7	5.7
Other	1,074	23	797	65	189	2.1	17.6
Genito-urinary:							
Polycystic kidney	122	9	97	6	10	7.4	38.5
External genitals	1,716	18	117	13	1,568	1.0	.1
Other	208	--------	130	15	63	--------	46.2
Bones and joints:							
Club foot	2,693	8	216	29	2,440	.3	--------
Chondrodystrophy	85	11	50	2	22	12.9	--------
Skull	957	45	183	11	718	4.7	2.3
Polydactyl	1,249	--------	75	3	1,171	--------	--------
Other	604	--------	43	5	556	--------	.5
Other malformations:							
Hemangioma	1,451	--------	74	18	1,359	--------	--------
Abdominal hernia	747	56	196	5	490	7.5	--------
Respiratory system	164	4	118	5	37	2.4	53.1
Webbed digits	310	1	11	2	296	.3	--------
Other skin	2,263	222	664	27	1,567	9.8	5.6
Mongolism	772	35	250	51	436	4.5	5.8

During the decade, there has been a consistent decline in the total malformation incidence rate from 1,428 to 1,144 cases per 100,000 births (fig. 1). An appreciable part of this reduction is due to a marked decline in the CNS malformations and club foot. A special effort was made to obtain data for the 1940 decade, and it was possible to construct incidence rates back to 1945 for this one group of defects (table 3). In the 15-year period between 1945 and 1960, each of the

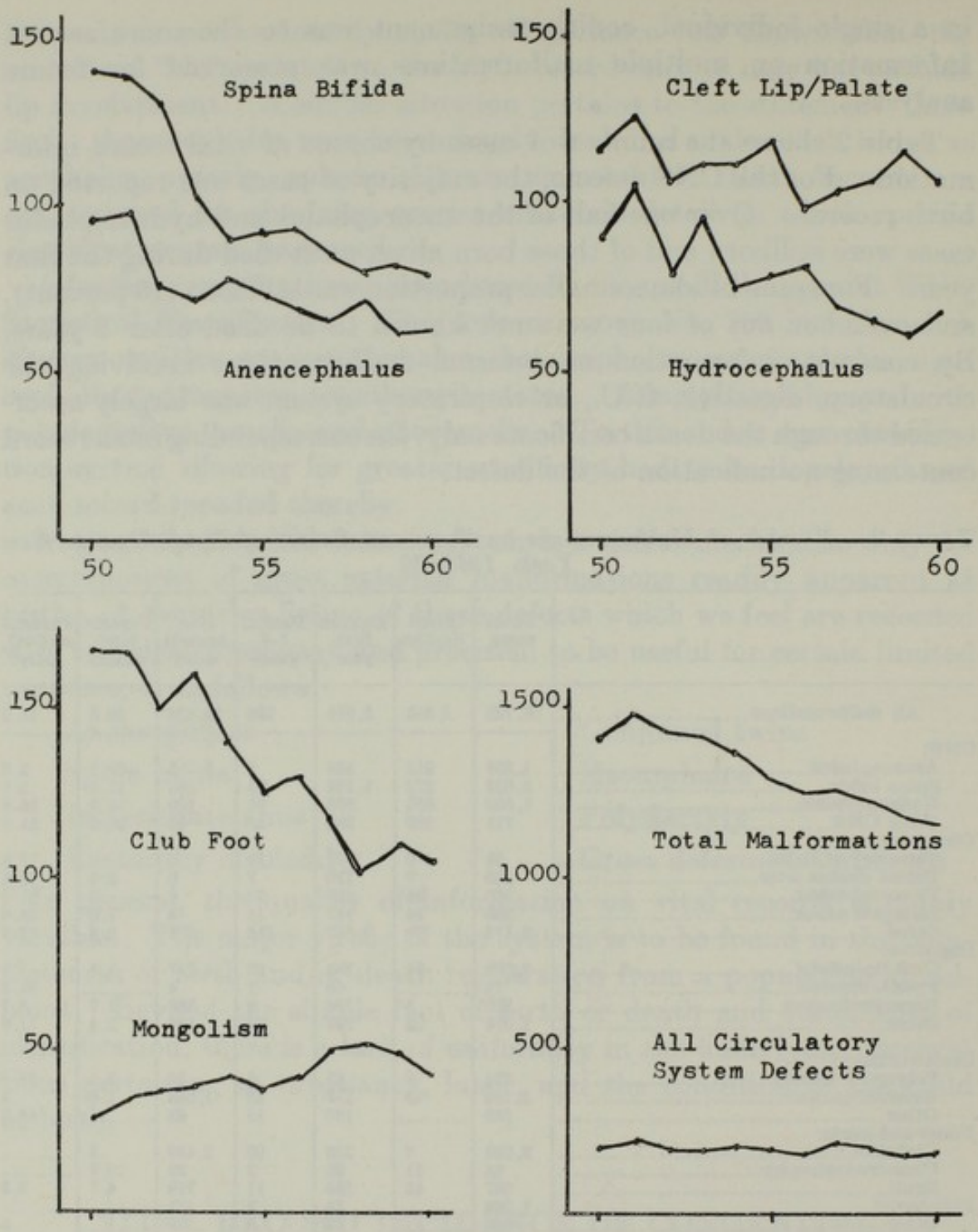


FIGURE 1.—Incidence of Selected Malformations by Year, Upstate New York, 1950–60 (Cases per 100,000 total births)

three major CNS malformations has decreased by as much as one-half. It is doubtful that reporting artefact could account for a downward secular trend of this magnitude, since it has occurred during a time of increasing awareness of, and interest in, the subject. During the same period the incidence of other defects, notably cleft palate and heart malformations, has remained fairly stable (table 4).

The declining rates for CNS and club foot defects make certain hypotheses being currently entertained less tenable. The 10- or 15-year period for which the trends have been measured is of short

TABLE 3.—*Incidence of Anencephalus, Spina Bifida, and Hydrocephalus by Year, Upstate New York, 1945–60*

Year	Total births (1,000's)	Cases per 100,000 births		
		Anencephalus	Spina bifida	Hydrocephalus
1945	110	120	192	141
1946	139	123	166	94
1947	159	116	165	103
1948	150	114	150	89
1949	152	117	156	92
1950	153	95	140	90
1951	164	98	138	105
1952	173	75	130	78
1953	175	72	101	96
1954	183	78	90	73
1955	190	73	92	77
1956	194	69	93	81
1957	203	65	84	68
1958	204	70	79	65
1959	204	60	83	60
1960	205	64	81	68
1945–60	2,758	83	117	83

TABLE 4.—*Number of Malformations by Type and Year, Upstate New York, 1950–60*

	Number of cases	Average 1950–60	Cases per 100,000 total births					
			1950–	1952–	1954–	1956–	1958–	1960
All malformations	26,728	1,305	1,453	1,423	1,333	1,256	1,196	1,174
CNS:								
Anencephalus	1,504	74	97	73	76	67	65	64
Spina bifida	2,025	99	139	116	91	88	81	81
Hydrocephalus	1,580	77	98	87	75	74	63	68
Other CNS	711	35	46	37	32	32	30	34
Circulatory	3,863	189	199	187	190	196	174	183
Digestive:								
Cleft lip/palate	2,259	110	121	109	116	100	110	105
Pyloric stenosis	71	4	6	5	5	3	2	2
Imperforate anus	300	15	16	17	16	11	15	18
Other G.I.	1,074	52	58	58	50	50	50	50
Genito-urinary:								
Polycystic kidney	122	6	5	6	7	5	6	8
External genitals	1,716	84	84	76	82	85	91	84
Other G.U.	208	10	8	8	10	12	13	9
Bones and joints:								
Club foot	2,693	132	166	155	132	124	106	105
Chrondrodystrophy	85	4	4	6	4	3	5	3
Malformations of skull	957	47	56	63	42	38	39	44
Digits	1,249	61	60	64	61	60	60	68
Other	604	29	36	32	33	27	25	27
Other malformations:								
Hemangioma and nevi	1,451	71	88	70	59	64	73	58
Hernia of abdominal cavity	747	36	39	28	36	39	38	38
Respiratory system	164	8	8	7	8	7	10	9
Webbed fingers and toes	310	15	17	13	16	14	13	19
Other skin	2,263	110	74	169	146	113	80	61
Mongolism	772	38	31	39	40	44	48	39
Total Births (1,000's)	2,047	186	317	348	373	398	408	205

enough duration as to make unlikely the operation of factors as predominant causes which remain fairly constant in the environment; *i.e.*, background radiation and certain meteorological variables. It is difficult to reconcile a simple genetic hypothesis with such short term changes. The pattern suggests that whatever the genetic predisposition to congenital defect may be, the phenotypic manifestations of that predisposition are very significantly influenced by varying

environmental factors, a point already clear from the well-known relationship between maternal age and frequency of defect.

Malformation rates are generally higher at the extremes of the maternal age range (figure 2, table 5a–5g). A similar pattern is noted for practically all malformations except polydactyly. In mongolism, the maternal age curve assumes its most striking form, empirical risk figures rising over tenfold with increasing age. Despite the better understanding of mongolism in terms of abnormal karyotype, the mechanism of the maternal age effect remains unclear.

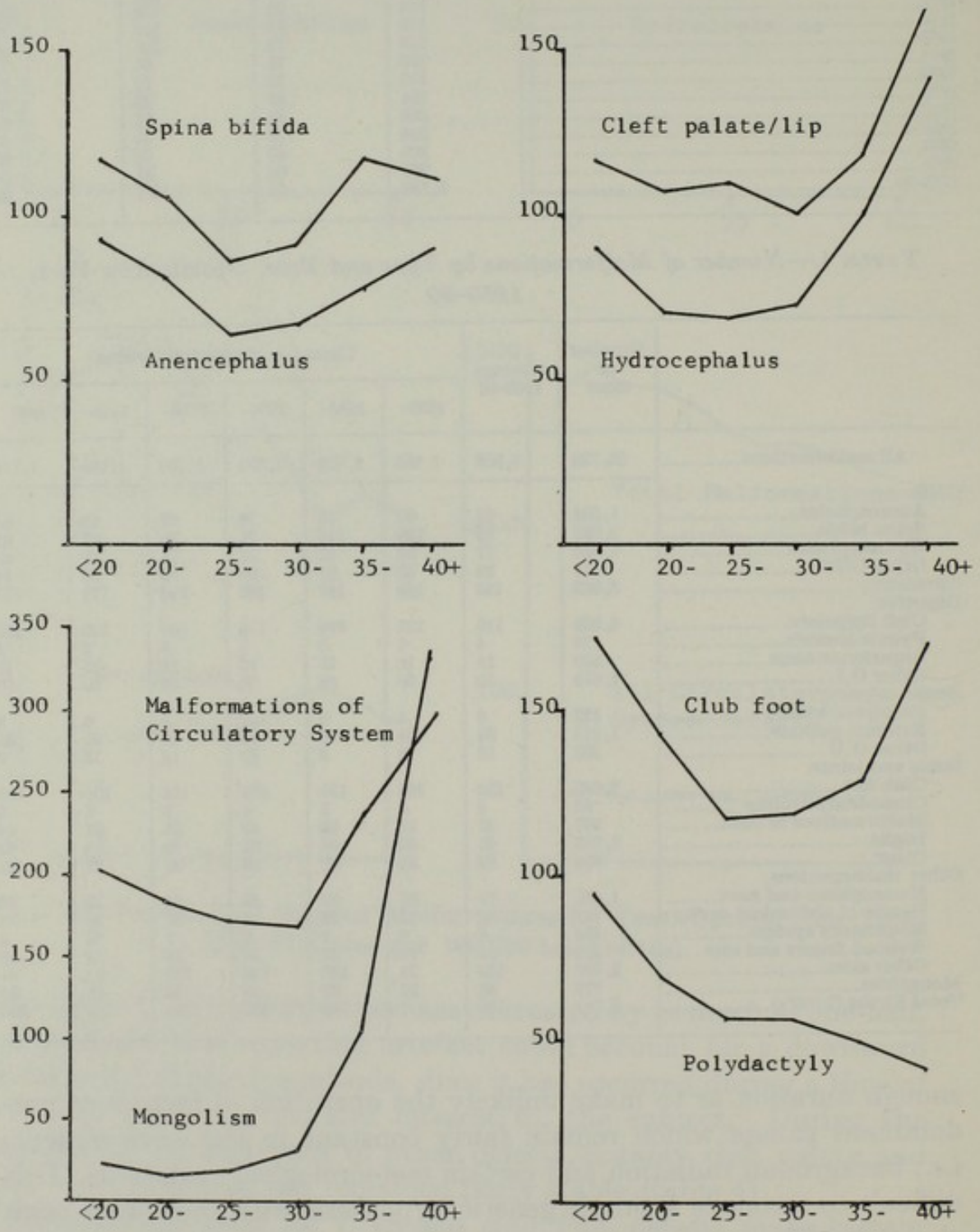


FIGURE 2.—Incidence of Selected Congenital Malformations by Maternal, Age, Upstate New York, 1950–60 (Cases per 100,000 total births)

TABLE 5a.—*Malformations by Type and Maternal Age, Upstate New York, 1950–60*

	Cases per 100,000 total births						
	Total	<20	20–24	25–29	30–34	35–39	40+
All malformations	1,305	1,527	1,325	1,184	1,198	1,492	2,107
CNS:							
Anencephalus	74	94	81	63	66	78	91
Spina bifida	99	116	106	86	92	118	107
Hydrocephalus	77	89	71	70	73	101	145
Other CNS	35	39	32	30	38	39	70
Circulatory	189	202	190	172	159	239	302
Digestive:							
Cleft lip/palate	110	117	108	110	101	119	171
Pyloric stenosis	4	7	4	2	3	4	6
Imperforate anus	15	17	16	14	13	16	16
Other G.U	53	60	55	43	47	64	113
Genito-urinary:							
Polycrystic kidney	6	5	5	7	6	5	8
External genitals	84	99	88	80	76	81	113
Other G. U	10	8	11	11	9	9	8
Bones and joints:							
Club foot	132	174	143	117	120	128	171
Chrondrodystrophy	4	2	4	5	2	7	14
Malformations of skull	47	66	51	42	44	45	42
Digits	61	95	68	56	55	49	40
Other	30	49	32	29	23	24	28
Other	241	269	244	229	227	257	330
Mongolism	38	20	17	19	33	108	334
Total births (1,000s)	2,047	147	592	631	424	202	50

TABLE 5b.—*Number of Cases of CNS Malformation by Maternal Age and Number of Previous Children, Upstate New York, 1950–60*

Type of malformation	Number previous children	Number of cases	Cases per 100,000 total births						
			Maternal age						
			Total	<20	20–24	25–29	30–34	35–39	40+
Anencephalus	Total	1,504	73.4	94.3	81.4	63.3	66.3	77.5	90.5
	0	474	85.9	94.2	90.3	67.2	87.5	98.6	80.6
	1	396	67.4	90.7	79.2	54.7	55.4	93.2	46.9
	2	263	60.9	115.6	57.2	63.3	56.7	55.6	107.0
	3	183	77.5	130.7	88.4	71.2	74.8	84.6	67.4
	4	85	73.4	-------	56.1	88.0	56.9	82.4	88.0
	5+	103	92.8	-------	188.3	58.7	96.8	83.3	149.7
Spina bifida	Total	2,025	98.9	116.0	105.5	86.0	92.5	118.0	106.5
	0	608	110.2	119.3	112.4	95.9	101.0	147.8	107.4
	1	541	92.1	100.8	98.0	81.8	88.6	116.5	124.9
	2	391	90.6	115.6	102.1	73.8	92.2	121.4	64.2
	3	226	95.7	130.7	91.7	103.7	89.0	92.1	89.9
	4	113	97.6	-------	157.1	90.9	79.2	121.6	73.4
	5+	146	131.6	(*)	125.5	88.0	121.7	148.0	193.7
Hydrocephalus	Total	1,580	77.2	88.9	71.1	69.7	73.4	100.8	144.7
	0	479	86.8	87.9	81.2	81.2	92.0	129.4	349.1
	1	361	61.5	87.3	56.5	56.1	52.4	110.7	156.2
	2	317	73.4	96.4	61.7	70.7	70.2	105.0	96.3
	3	199	84.2	261.4	111.3	82.4	73.5	82.2	78.6
	4	94	81.2	-------	67.3	70.4	94.0	58.8	161.4
	5+	130	117.1	-------	94.1	79.6	99.3	141.9	193.7
Other CNS malformations	Total	711	34.7	39.3	31.9	29.8	38.0	39.0	70.4
	0	196	35.5	40.4	31.5	28.7	40.4	80.1	107.4
	1	181	30.8	33.6	33.8	25.7	30.2	40.8	46.9
	2	136	31.5	38.5	28.0	29.8	31.3	39.1	53.5
	3	80	33.9	-------	26.2	35.0	25.8	37.3	101.1
	4	57	49.2	-------	44.9	38.1	64.3	27.5	102.7
	5+	61	55.0	-------	62.8	41.9	74.5	33.9	61.6
Total births (thousands)	----------	----------	2,047	147	592	631	424	202	50

735–712 O—65——21

TABLE 5c.—*Number of Cases of Hare Lip and Cleft Palate Malformation by Maternal Age and Number of Previous Children, Upstate New York, 1950–60*

Type of malformation	Number previous children	Number of cases	Cases per 100,000 total births						
			Maternal age						
			Total	<20	20–24	25–29	30–34	35–39	40+
Hare lip and cleft palate	Total	2,259	110.3	116.7	107.9	110.4	100.6	118.5	170.9
	0	571	103.5	121.1	96.6	104.1	87.5	123.2	134.3
	1	648	110.3	110.9	120.2	101.3	98.7	139.8	109.3
	2	478	110.7	77.1	113.3	119.1	96.4	102.9	181.9
	3	266	112.6	----	101.5	118.7	103.2	112.0	168.5
	4	129	111.4	----	78.6	102.7	99.0	129.4	205.4
	5+	167	150.5	----	188.3	146.7	136.6	135.7	237.7
Total births (thousands)	----	----	2,047	147	592	631	424	202	50

TABLE 5d.—*Number of Cases of Genito-Urinary Malformation by Maternal Age and Number of Previous Children, Upstate New York, 1950–60*

Type of malformation	Number previous children	Number of cases	Cases per 100,000 total births						
			Maternal age						
			Total	<20	20–24	25–29	30–34	35–39	40+
Genito-urinary	Total	1,716	83.8	99.0	87.9	80.3	76.0	81.0	112.6
	0	561	101.7	105.0	99.4	96.8	96.5	129.4	268.5
	1	459	78.2	84.0	82.5	75.1	72.5	69.9	140.6
	2	329	76.2	77.1	70.7	75.7	70.2	102.9	74.9
	3	179	75.8	----	81.9	88.7	65.8	69.7	44.9
	4	85	73.4	----	67.3	64.5	81.6	62.7	117.4
	5+	103	92.8	----	125.5	67.1	99.3	77.1	158.8
Total births (thousands)	----	----	2,047	147	592	631	424	202	50

TABLE 5e.—*Number of Cases of Club Foot Malformation by Maternal Age and Number of Previous Children, Upstate New York, 1950–60*

Type of malformation	Number previous children	Number of cases	Cases per 100,000 total births						
			Maternal age						
			Total	<20	20–24	25–29	30–34	35–39	40+
Club foot	Total	2,693	131.5	173.7	143.2	116.9	119.5	128.4	170.9
	0	984	178.4	192.0	174.3	166.5	188.5	197.1	241.7
	1	704	119.9	120.9	125.0	114.6	112.8	125.2	203.0
	2	462	107.0	96.4	113.3	101.7	102.3	121.4	128.4
	3	262	110.9	130.7	91.7	96.2	112.2	139.4	146.0
	4	127	109.6	----	190.8	90.9	94.0	117.6	161.4
	5+	154	138.8	----	31.3	92.2	158.9	123.4	237.7
Total births (thousands)	----	----	2,047	147	592	631	424	202	50

TABLE 5f.—*Number of Cases of Adactylism and Polydactylism Malformation by Maternal Age and Number of Previous Children, Upstate New York, 1950–60*

Type of malformation	Number previous children	Number of cases	Cases per 100,000 total births						
			Maternal age						
			Total	<20	20–24	25–29	30–34	35–39	40+
Adactylism, polydactylism.	Total	1,249	61.0	95.0	67.7	56.2	55.0	49.4	40.2
	0	329	59.6	78.1	54.0	54.1	62.8	61.6	26.9
	1	328	55.8	131.0	61.8	47.1	43.3	40.8	78.1
	2	254	58.8	231.3	81.9	49.0	53.3	49.4	32.1
	3	132	55.9	130.7	117.9	56.2	40.0	37.3	44.9
	4	95	82.0	--------	269.3	93.9	61.9	47.1	29.3
	5+	111	100.0	--------	94.1	142.5	106.8	77.1	44.0
Total births (thousands)	----------	----------	2,047	147	592	631	424	202	50

TABLE 5g.—*Number of Cases of Mongolism by Maternal Age and Number of Previous Children, Upstate New York, 1950–60*

Type of malformation	Number previous children	Number of cases	Cases per 100,000 total births						
			Maternal age						
			Total	<20	20–24	25–29	30–34	35–39	40+
Mongolism	Total	772	37.7	20.4	16.7	18.8	33.1	107.7	333.7
	0	135	24.5	26.0	15.0	18.0	40.4	110.9	268.5
	1	176	30.0	3.4	15.9	17.1	35.2	107.8	531.0
	2	185	42.8	--------	20.2	22.3	38.1	137.9	203.3
	3	119	50.4	--------	29.5	17.5	27.1	102.1	381.9
	4	71	61.3	--------	11.2	23.5	27.2	98.0	381.5
	5+	86	77.5	--------	--------	12.6	24.8	92.5	378.6
Total births (thousands)	----------	----------	2,047	147	592	631	424	202	50

The relationship between malformation rate and birth order is similar to the maternal ageing effect, the risk being elevated for primiparas and for the higher parities. A joint analysis reveals a birth order effect which is independent of maternal age. In hydrocephalus, elderly primiparas have rates far higher than elderly multiparas.

Malformation rates by type, sex, and race, are presented in table 6. The malformation rate in males is higher than that in females, with certain well known exceptions. (Anencephalus and spina bifida show a female excess.) Unfortunately, cleft palate is in the same code category as hare lip, and the female excess expected in isolated cleft palate is, therefore, obscured. The tremendous male excess in malformations of external genitalia is due to the presence of hypospadias and epispadias in this category.

The most striking racial differences are seen in the well known nonwhite excess of adactyly and polydactyly. If these malformations are not considered, the white total malformation rate will be considerably higher than the nonwhite. The white excess is quite apparent in anencephalus and spina bifida. In anencephalus, the excess is

TABLE 6. *Malformation Cases by Type, Sex, and Color, Upstate New York, 1950–60*

Type of malformation	Number of cases				Cases per 100,000 total births			
	White		Nonwhite		White		Nonwhite	
	M	F	M	F	M	F	M	F
All malformations	14,238	10,928	818	696	1,419.8	1,150.6	1,688.1	1,516.5
CNS:								
Anencephalus	492	970	21	17	49.1	102.1	43.3	37.0
Spina bifida	850	1,127	20	22	84.8	118.7	41.3	47.9
Hydrocephalus	853	654	34	36	85.1	68.9	70.2	78.4
Other CNS	296	369	20	18	29.5	38.9	41.3	39.2
Circulatory	2,160	1,506	92	105	215.5	158.5	189.9	228.8
Digestive:								
Cleft lip/palate	1,316	884	33	25	131.2	93.1	68.1	54.5
Pyloric stenosis	52	14	2	3	5.2	1.5	4.1	6.5
Imperforate anus	215	73	5	4	21.4	7.7	10.3	8.7
Other G.I.	594	411	33	36	59.2	43.3	68.1	78.4
Genito-urinary:								
Polycystic kidney	81	35	4	2	8.1	3.7	8.3	4.4
External genitals	1,581	73	48	7	157.7	7.7	99.1	15.3
Other G.U.	148	51	7	2	14.8	5.4	14.4	4.4
Bones and joints:								
Club foot	1,456	1,137	56	44	145.2	119.7	115.6	95.9
Chondrodystrophy	32	46	3	4	3.2	4.8	6.2	8.7
Malformations of skull	458	424	38	36	45.7	44.6	78.4	78.4
Digits	508	317	247	177	50.7	33.4	509.7	385.7
Other	250	337	5	10	24.9	35.5	10.3	21.8
Other malformations:								
Hemangioma and nevi	707	673	36	35	70.5	70.9	74.3	76.3
Hernia of abdominal cavity	443	222	30	51	44.2	23.4	61.9	111.1
Respiratory system	94	64	2	4	9.4	6.7	4.1	8.7
Webbed fingers and toes	190	110	6	4	18.9	11.6	12.4	8.7
Other skin	1,119	1,011	74	47	111.6	106.4	152.7	102.4
Mongolism	343	420	2	7	34.2	44.2	4.1	15.3

limited to white females, while in spina bifida, white males and females have rates twice those of nonwhites. The nonwhite excess in hernia of the abdominal cavity is probably part of the continuum which finds umbilical hernia more common in nonwhite than white infants.

CONGENITAL MALFORMATION IN TWINS

Clearly, a case by case study of malformation incidence can furnish few clues in evaluating the relative roles of heredity and environment as etiologic factors. Outstanding exceptions include the relationships between rubella and the "rubella" syndrome, and between thalidomide and phocomelia. The possibility of developing sibship information through vital record linkage represents one avenue of approach. One sibship analysis has already been completed, the probands being all cases of anencephalus or spina bifida occurring within one city of the State (Milham, 1962).

Along a similar line, twins may be regarded as partial sibships. Further, the two distinct twinning processes of polyovulation and polyembryony provide a naturally occurring mechanism which simulates controlled experimentation. The difference in origin of the two types of twins has significant genetic implications for, while members

of dizygotic sets may vary at given genic loci, members of monozygotic sets are of identical genotypes. The utility of the twin method is based on this difference. A genetically determined trait should be reflected in a higher concordance among monozygotes than dizygotes.

All twins among the 2 million births occurring during the study period were matched with their corresponding mates to form a twin maternity. The total number of such sets was 21,128. Unfortunately, zygosity classification is not available on the birth record and indirect methods must be relied upon. Present knowledge would suggest that male-female sets are dizygotic in that an entire chromosome ordinarily differentiates the pair. By contrast, like sex sets may be of either zygosity type. Intersexes, supersexes and other conditions arising out of nondisjunction, translocation and failure of segregation of the sex chromosomes, are too infrequent to upset seriously the general pattern. Certainly, it would be possible for uniovular twins to be of opposite sex—one an XO "female" with

TABLE 7. *Twin Sets by Sex Composition and Number Affected in Set for Selected Characteristics, Upstate New York, 1950–60*

Characteristic	Number affected in set				Cases per 100,000		
	Like sex		Unlike sex		MM, FF twins	MF, FM twins	Single births
	One	Both	One	Both			
Maleness	0	7,300	6,818	0	(*)	(*)	(*)
Femaleness	0	7,010	6,818	0	(*)	(*)	(*)
Twinness	0	14,310	0	6,818	(*)	(*)	(*)
"Siamese" trait	0	12	0	0	84	0	0
CNS malformations	102	12	49	2	430	359	285
Anencephalus	24	4	14	1	112	117	74
Spina bifida	24	3	15	1	105	125	99
Hydrocephalus	33	1	15	0	122	110	77
Eye Agenesis	1	1	1	0	10	7	5
Aplasia of brain	7	2	0	0	38	0	9
Monster, NOS	13	1	4	0	52	29	24
Circulatory system	70	6	27	0	286	198	189
Genito-urinary system	27	5	12	0	129	88	100
Gastro-intestinal system	49	6	29	0	199	212	181
Cleft lip/palate	31	5	14	0	143	103	110
Pyloric stenosis	4	0	1	0	14	7	4
Imperforate anus	8	0	4	0	28	29	15
Other malformations	102	14	71	1	454	535	513
Club foot	38	2	26	0	147	191	132
Polydactyly	22	4	8	0	105	59	61
Mongolism	6	1	5	0	28	37	38

*The following pairs, each individual with a different malformation, were treated as discordant:

Like sex sets

1 anencephalus, spina bifida
1 anencephalus, hydrocephalus
2 monster nos, imperforate anus
1 monster nos, polydactyly
1 circulatory, cleft palate
1 circulatory, TE fistula
1 G.U., polydactyly
1 G.U., other
1 Other, mongol

Unlike sex sets

1 spina bifida, hydrocephalus
1 monster nos, mongol
1 cleft palate, bone and joint

Turner's syndrome and its mate, an XXY "male" with Klinefelter's syndrome.

Using sex concordance and sex discordance as an index of zygosity, table 5f shows the recorded malformation cases among twin sets by the number affected within each set. In all, there were 14,310 like sex sets, of which 7,300 sets were both male, and 7,010 both female; and 6,818 sets, of which one was male and the other female. A total of 12 pairs, all among the like sex sets, were conjoined (Siamese trait). Confidence in the data would have been reduced had it been otherwise, since podencephaly, syncephaly, dicephalus dipus dibrachius, and the like, presumably are manifestations of incomplete monozygotic twinning.

Excluding conjoined twins, a condition lumped with "monstrosity" in the International List, a total of 409 of the like sex sets and 189 of the unlike sex sets had one or two members with malformations noted on their vital records. The total malformation rate among twin births was 1.5 percent, which is only slightly in excess of the rate of 1.3 percent for all births in the population.

A comparison of the number affected for given defects within like and unlike sex sets reveals that 44 of the 409 sets among the former were concordant, while only 3 of the 189 sets among the latter were concordant. The average concordance coefficient for all malformations in like sex sets was r=.18 while for unlike sex sets, r=.01 (see Gittelsohn and Milham, 1963). This suggests that monozygotes have a higher concordance than dizygotes and an average heritability of about 29 percent for all congenital malformations.

For major malformations of the central nervous system, observed numbers affected among like and unlike sex sets and estimated numbers affected among monozygotic and dizygotic sets were as follows:

Number affected		Like sex	Unlike sex	Mz	Dz
Total sets		14,310	6,818	7,492	13,636
0	Concordant	14,196	6,767	7,429	13,534
1	Discordant	102	49	53	98
2	Concordant	12	2	10	4
Rate per 100,000 births		430	359	428	359
Concordance coefficient (R)		.19	.04	.26	.04

Sets with each member affected with different CNS defects were classified as discordant. The Weinberg assumption was used for zygosity estimation. With the usual definition of the "Heritability Index"

$$H=1-R(MZ)/R(DZ)$$

we find that H=23 percent.

Several problems present themselves in such an analysis. A sex-selective incidence for a given trait would tend to reduce the concordance for that trait in unlike sex sets. At the extreme, such characteristics as hypospadias, bicornate uterus, and ovarian agenesis will be completely discordant in male-female sets. Anencephalus with a 2:1 female predominance, and imperforate anus with a 3:1 male predominance, will behave in a similar fashion. Subsetting by MM and FF among the like sex will only clarify the matter in part, at the expense of reducing the numbers available for analysis in any one cell. Even with a direct zygosity diagnosis, the usual limitations of the twin method remain extant.

CONCLUSION

Vital records can be a valuable source of congenital malformation information. Since malformations are comparatively rare events, it is difficult for an individual or a hospital to collect case series large enough for meaningful analysis. The official agency which collects vital records, however, is in a uniquely advantageous position inasmuch as it collects data on nearly all births, for which in the case of malformations, the entire population is at risk.

Vital records have certain limitations as a source of malformation information. Recognizing these limitations and working within them, data may be obtained which are elsewhere unobtainable.

SUMMARY

1. Malformations reported on vital records in the years 1950–60, in upper New York State, have been tabulated and analyzed.

2. The total malformation rate is decreasing, with central nervous system malformations showing the most marked reduction.

3. Incidence rates by sex, race, maternal age, and birth order have been presented.

4. The concordance of malformations is higher in like sex than in unlike sex twin sets, suggesting a zygosity differential.

REFERENCES

1 Gittelsohn, A. and Milham, S. (1962). The declining incidence of central nervous system anomalies in New York State. Brit. J. Prev. Soc. Med. 16: 153–158.

2. Gittelsohn, A. and Milham, S. (1963). Statistical study of twins. Am. J. Pub. Health. (In press.)

3. Milham, S. (1962). Increased incidence of anencephalus and spina bifida in siblings of affected cases. Science 138: 593–594.

4. Milham, S. (1963). Underreporting of incidence of cleft lip and palate. Am. J. Dis. Child. (In press.)

Methodological Problems in the Use of Standard Vital Statistical Data in the Study of Neonatal Mortality and Birth Weight*

By Douglas Grahn, Ph. D., *Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Ill.*

A recent report by Grahn and Kratchman (1963) concerned itself with the evaluation of neonatal death rate and birth weight in selected regions of the United States, with particular reference to the possible effects of the geologic environment, radiation and altitude. The results of that study indicate the neonatal death rate to be high in the western mountain states of Arizona, New Mexico, Colorado, Nevada, Wyoming, the western portions of Nebraska, South Dakota, North Dakota, Kansas, and the Texas Panhandle, as compared to Illinois and Indiana or the U.S. average. A significant positive relationship was shown to exist between death rate and altitude. The increased death rate was largely attributed to a lower birth weight, since the frequency of immature births, on a weight criterion, progressively increases with altitude. The observed correlations with altitude were evaluated in terms of either the increase in cosmic radiation intensity or the decrease in oxygen partial pressure, or both. The data were evaluated in terms of three hypotheses: radiation-induced mutation, radiation-induced injury to the fetus, and hypoxia-induced depression of fetal growth. Very little of the excess neonatal death could be attributed to readily-defined genetic factors, and direct fetal irradiation was concluded to be of no significance. The weight of the evidence—historical, experimental and clinical—strongly suggested that the reduced partial pressure of oxygen is responsible for a reduced fetal growth rate and increased neonatal death rate.

The original purpose of the study was to detect the existence of any possible relationship between early mortality and environmental

*This work was performed under the auspices of the U.S. Atomic Energy Commission.

radiation with particular reference to the major uranium-ore-bearing regions of the U.S. west of the Mississippi. Interest in this problem has increased in recent years due to public concern over the level of radiation added to man's environment as a result of the testing of nuclear weapons. With or without this source of low-level radiation, the question requires evaluation to prepare for the progressive increase in the application of nuclear energy in the form of stationary civilian power reactors, nuclear rockets, and marine propulsion reactors as well as medical and agricultural uses. Both genetic and somatic endpoints can be studied with equal pertinence to the problem, but genetic endpoints are preferred because they are not confounded by a lifetime's accumulation of environmental experiences.

This brief discussion of methodological problems will be restricted to the use of vital statistical and census data in studies of radiation epidemiology. The data from the Grahn-Kratchman (1963) study will be exploited for this purpose. Direct field evaluation is certainly preferred, but a broad sweep across the continent, for example, through the use of the available vital statistical data, can be extremely profitable in detecting areas of potential value for further study. In addition, in the material used for this discussion, this approach clearly defined an environmental problem of unrestricted scope.

SOURCES OF DATA

The vital statistical data were obtained directly from the tabulations published annually by the U.S. Department of Health, Education and Welfare (1950–1957). The specific counties and states involved are given in Grahn and Kratchman (1963). Tabulations were made of the number of neonatal deaths and the number of live births by county of residence for the white population. Birth weight data were tabulated as the percentage of live births at "2,500 grams or less," the recognized category of births that is classed as "immature" in the Sixth Revision of the International Lists of Diseases and Causes of Death.

Several additional tabulations were made for selected states, regions or years. These include: gestation length, mean birth weight, percent born in hospitals, cause of death, and time of death. For the states of Delaware, Illinois, Indiana, Idaho and Montana, which contain virtually no uranium ore reserves, neonatal mortality rates were tabulated on a statewide basis rather than by summation of individual counties for the purpose of rough comparison across an array of geologic environments.

Certain additional characteristics of the surveyed population groups were tabulated to evaluate differences in age and socio-economic level. Median family income and median age values were tabulated, by county, from the published 1950 U.S. census data

(U.S. Bureau of the Census, 1953). Where needed, other characteristics, such as median school years completed, percent nonwhite, total population number, and population growth figures were also derived from the census reports.

Two tables of data from the Grahn-Kratchman study are presented here to orient the reader to the general problem and the final data. Table 1 gives the age- and income-adjusted neonatal death rates by State along with the associated factors of altitude, atmospheric pressure, annual cosmic radiation dosage and uranium ore tonnage. Table 2 presents the unadjusted birth-weight data by class intervals of 500 or 1,000 feet of altitude from sea level to 11,000 feet. The data for any interval are based upon the county altitude values without regard to State or geologic province for all states except Utah.

TABLE 1.—*Age and Income Adjusted Neonatal Death Rates and Associated Physical Variables by State: 1950–57, White Population Only*

State	Altitude (feet)	U_3O_8 reserves (tons)	Atmospheric pressure (mm Hg)	Cosmic radiation (mr/yr)	Number of live births	Neonatal deaths ±SE (per 1,000 births)
Texas (Coast)	310	570	752	35.1	37,216	18.13±0.98
Washington	1,440	(*)	721	41.9	11,078	18.01±1.80
California	1,490	40	720	42.1	26,203	18.33±1.17
Arizona	2,080	2,240	703	46.4	179,659	20.67±0.47
Kansas	3,080	(Helium)	678	53.6	15,007	20.90±1.65
Dakotas	3,180	2,680	676	54.2	33,975	20.10±1.08
Texas (Panhandle)	3,370	(Helium)	672	55.3	35,192	21.26±1.09
Nevada	3,690	190	664	58.0	40,360	22.64±1.04
Nebraska	3,920	----------	657	60.2	22,790	21.72±1.37
New Mexico	5,010	80,330	632	69.6	189,242	23.41±0.49
Colorado	5,390	11,670	623	73.1	306,818	22.19±0.38
Wyoming	5,390	60,830	623	73.1	66,728	22.65±0.82
Delaware	60	----------	758	34.0	63,648	17.10±0.73
Illinois and Indiana	630	----------	744	36.8	2,285,385	17.40±0.12
Idaho	3,320	----------	673	55.1	131,306	18.25±0.52
Montana	3,490	50	668	56.7	127,660	18.66±0.54
Utah	4,590	12,190	642	65.6	191,042	16.63±0.41

* Data not available for publication.

TABLE 2.—*Percentage of Live Births at 2,500 Grams or Less by Altitude Interval: 1952–57, White Population Only. Utah Not Included*

Altitude interval (feet)	Mean altitude	Atmospheric pressure (mm Hg)	Number of live births	Percent 2,500 grams or less
0–500	263	753	35,166	6.57
501–1,000	633	743	7,147	6.66
1,001–1,500	1,118	729	73,318	6.17
1,501–2,000	1,786	713	13,809	7.97
2,001–2,500	2,286	699	56,570	7.78
2,501–3,000	2,864	684	12,207	7.14
3,001–3,500	3,256	674	50,933	8.24
3,501–4,000	3,756	662	60,226	8.46
4,001–4,500	4,287	649	63,160	8.67
4,501–5,000	4,824	636	100,420	9.47
5,001–5,500	5,237	627	133,617	10.37
5,501–6,000	5,661	617	28,011	9.80
6,001–6,500	6,149	605	53,899	10.74
6,501–7,000	6,767	519	26,619	11.54
7,001–7,500	7,213	582	10,712	11.17
7,501–8,000	7,721	570	9,427	13.04
8,001–9,000	8,519	553	2,474	12.93
9,001–10,000	9,568	532	887	16.57
10,001–11,000	10,410	513	1,697	23.7

METHODS OF DATA REDUCTION

NEONATAL MORTALITY

The neonatal mortality data for 260 counties for the 8 years, 1950 to 1957, represent a formidable task for data reduction since the observed data vary from values near zero to over 50/1000. Sample sizes range from a few dozen to over 10,000 per year per county, but most samples are large enough to lead to modest confidence in the observed figures. Some of the more obvious sources of variation are itemized and discussed in the following.

A. *Completeness and accuracy of registration*

The introductory remarks of the annual vital statistics reports discuss the question of incomplete registration of births and deaths. Birth registration was about 99 percent complete for the white population in 1956. Death records are probably not as complete and vary in the validity and accuracy of the reported causes of death. No truly quantitative measure of registration completeness is available and diagnostic accuracy cannot be ascertained since generally less than 20 percent of all deaths are subject to autopsy. Data obtained since 1950 are more reliable than earlier data.

B. *Population base*

Neonatal mortality rates are customarily calculated in terms of the number of deaths in the first 28 days of life per 1,000 live births. The actual number of live births at risk will continually vary, even in a stable population, but the neonatal period is sufficiently short so that the summarizing procedure by yearly intervals will act to smooth out all fluctuations except those that might result from a sudden precipitous rise or fall in the birth rate or rate of migration.

C. *Calendar year*

During the 8-year study period, neonatal mortality declined 15 percent and the rate of change was similar for the mountain States as a group compared to the U.S. average. The mountain States, however, maintained a constant deviation of about 4 deaths per 1,000 above the U.S. average. Prior to 1950 the difference may have been greater, as judged from the total infant mortality data. (Neonatal deaths are not separately available before about 1950.) Figure 1 presents the data for selected States and attention is drawn to Arizona and New Mexico. These two States have reduced infant mortality by a factor of three to four while the U.S. average dropped by less than a factor of two. When genetic interpretations are to be examined, differential mortality rates as those indicated in figure 1 can create confusion, since changing public health practices will be ac-

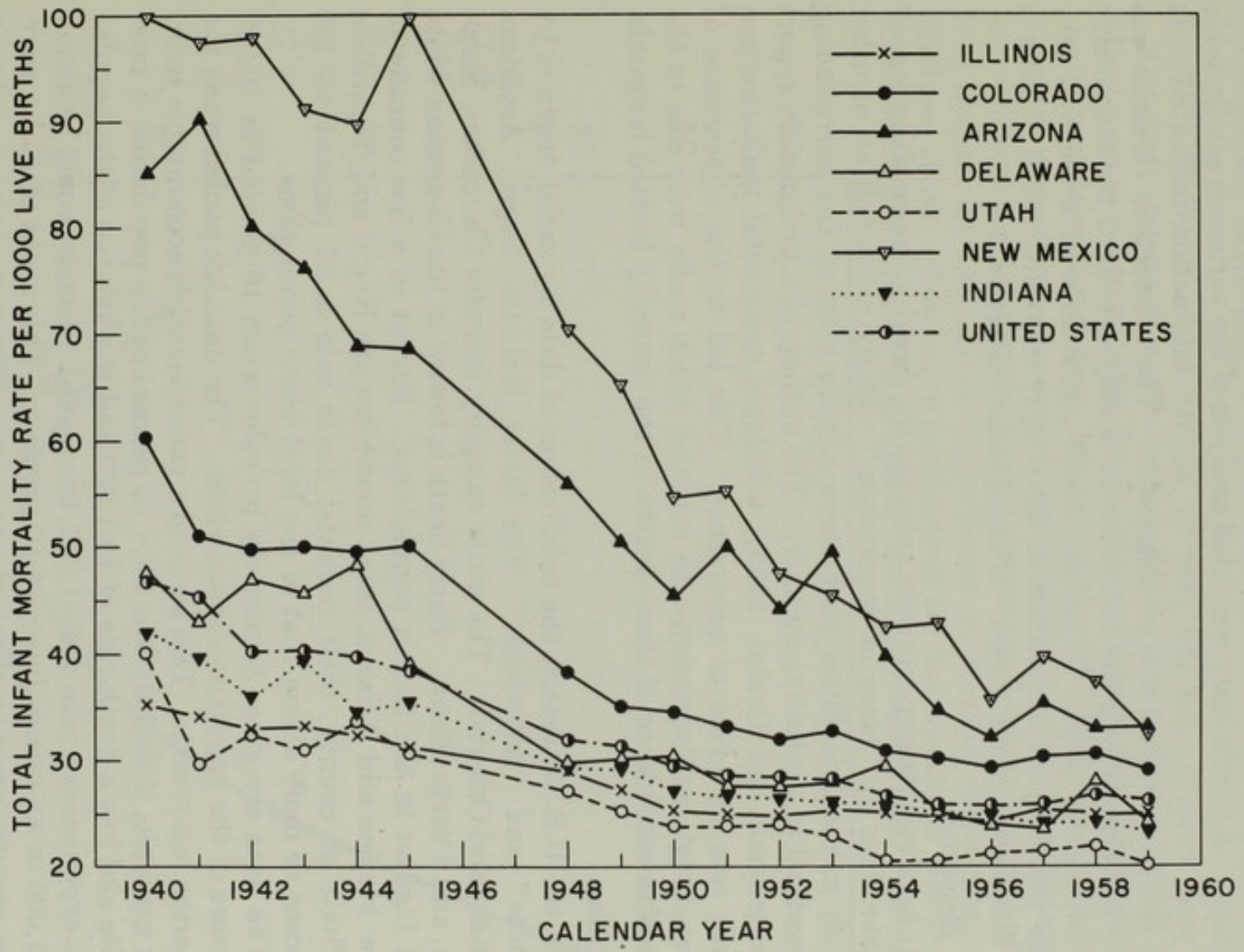


FIGURE 1.—Infant mortality data for indicated States. White population only.

companied by a changing selection pressure and probably as well by a changing accuracy of diagnosis of the causes of death. Studies requiring data from periods prior to 1950 would certainly suffer more than those involving recent years.

In the absence of good cause to examine the data on a year-to-year basis, the data for the study under discussion were pooled across the 8 years. A preliminary weighted analysis of the variances attributable to year, county, and the year by county interaction indicated that the county variance was predominant. The interaction variance was very uniform among the States and was only one-third the magnitude of the county term. Thus, one could conclude that there was very little year-to-year fluctuation in the county mortality rate. This will be discussed further under "*Unassignable Variation.*"

D. *Residence*

It is extremely important to tabulate the data by "place of residence" rather than by "place of occurrence." Many metropolitan or other urban medical facilities may serve a wide area of the surrounding hinterland in a selective manner. The author, in a preliminary report (Kratchman and Grahn, 1959), tabulated congenital malformation death rate by place of occurrence. This led to the appearance of several high malformation-rate counties which really were due to the selective occurrence of these deaths in the centrally located hospitals.

E. *Race*

In the U.S. statistics, the subdivision of data by racial origin is by "white" and "nonwhite." The latter includes Negro, American Indian, and Oriental. The white category includes Caucasian, Mexican, and Puerto Rican. Early death is lowest in the Caucasian white and highest in the Indian populations. Except in a few counties of New Mexico and Arizona, most nonwhites are Negro and, regardless of State or county, their neonatal death rate is 30 percent to 100 percent or more above that of the local white populations.

The data are given separately by color when 10 percent or 10,000 or more of the population is nonwhite. The data are pooled when the criterion is not met. This leads to a variable and unaccountable error. For example, if slightly less than 10 percent of the population is nonwhite and has a death rate of 40/1,000 and the balance of the population experiences a rate of 20/1,000, the pooled estimate would be about 22/1,000 or an overestimate of 10 percent. This could often be more than enough error to eliminate any true differences being sought or to create artificial differences.

An additional bias for which there is no present control involves the Mexican or "Spanish-American" populations of the southwestern United States. Judging from census figures on the number of foreign-born whites of Mexican origin, the inclusion of this group in some counties produces a serious elevation of the neonatal death rate in the

white population. Certain counties of Arizona, New Mexico, and Texas are particularly affected. A separate supplementary analysis of the data from Imperial County, Calif. (on the southern border) which has an average altitude below sea level, could not contribute to the main analysis of altitude relationships because of this uncontrollable heterogeneity of the data.

F. *Maternal age*

Although the use of death rates in terms of live births avoids age-adjustment requirements for the population base, it does not eliminate the maternal age factor from influencing the probability of death itself. Figure 2 presents the maternal age: neonatal mortality relationship derived from a special study carried out on the registered births of the first quarter of 1950 (U.S. Department of Health, Education, and Welfare, 1958). The curve is described by the equation: $Y=49.8-2.33A+0.042A^2$, where Y is mortality and A is maternal age. Although

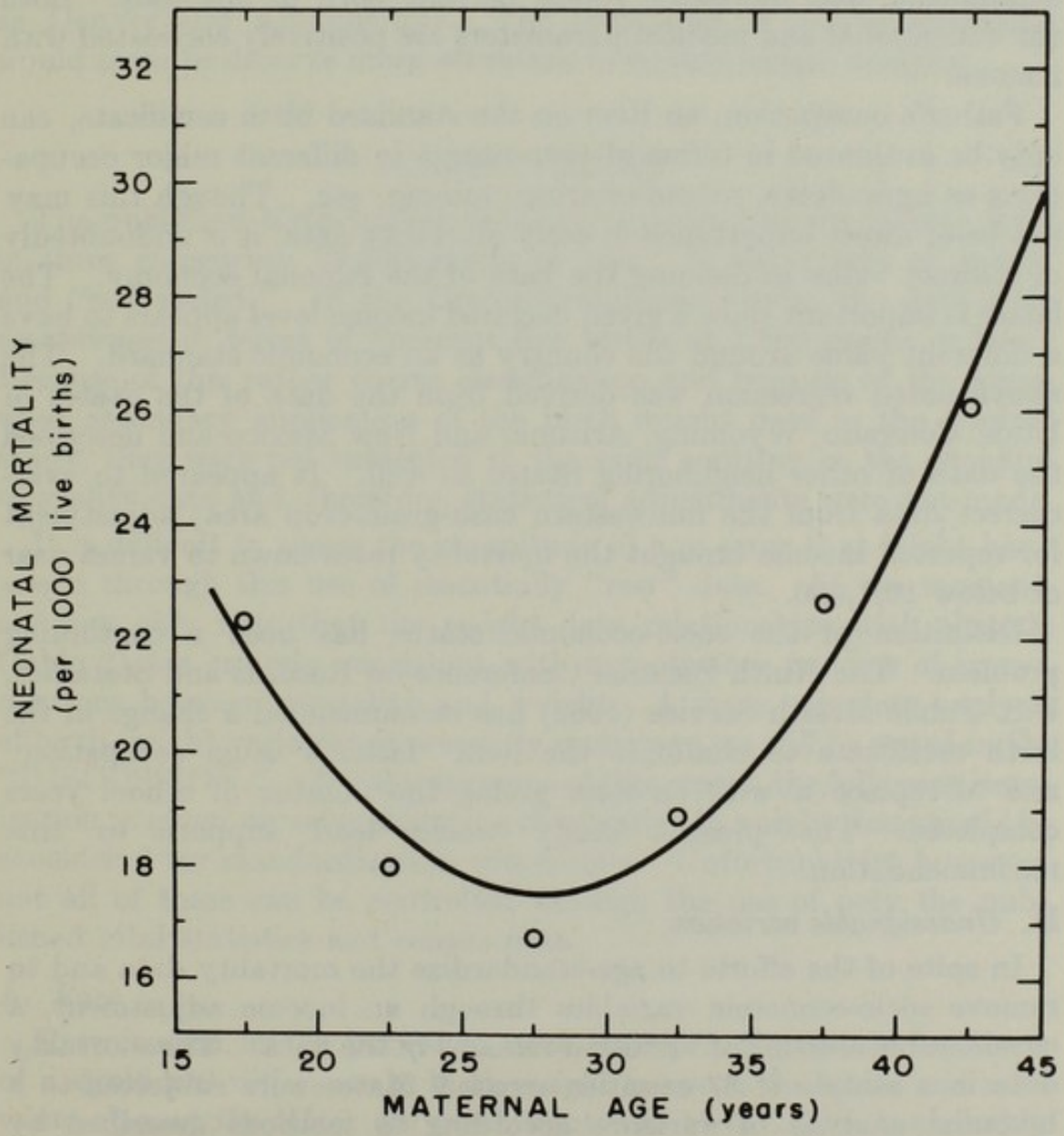


FIGURE 2.—Regression of neonatal death rate on maternal age, 1950, white population.

this equation was used to adjust all county and whole-State values to a constant maternal age in the Grahn-Kratchman study, there is no way of ascertaining its validity for all regions of the U.S. Age must be accounted for since it varies from 20 to 40 years among the counties in the Mountain States.

G. *Socio-economic factors*

In the Grahn-Kratchman study, an analysis of the interrelationships among the following variables was carried out to evaluate any effect on neonatal mortality: median income, median age, median number of school years completed, percent born in hospitals. It was concluded that the median income best defined the socioeconomic status and county mortality rates were adjusted to a constant income by means of a linear regression of mortality rate on income of −2.0/1,000 live births per $1,000 median income. Good negative correlations were also noted, however, between mortality rate and educational attainment, and mortality versus percent born in hospitals. Both the educational and medical parameters are positively correlated with income.

Father's occupation, an item on the standard birth certificate, can only be evaluated in terms of percentages in different major occupations as agriculture, manufacturing, mining, etc. Though this may not be of direct importance in early mortality data, it is undoubtedly of indirect value in defining the base of the regional economy. The latter is important since a given declared income level appears to have a different value around the country as an economic standard. The above noted regression was derived from the data of the States of Utah, Colorado, Wyoming, Arizona, and New Mexico and described the data of other neighboring States as well. It appeared to overcorrect data from the midwestern cash-grain-crop area; adjustment for reported income brought the mortality rates down to values near or below 10/1,000.

Definition of the socio-economic status has been a continuing problem. The Ninth Biennial Conference on Records and Statistics, U.S. Public Health Service (1962) has recommended a change in the birth certificate to eliminate the item "father's usual occupation" and to replace it with an item giving the number of school years completed. The present study would lend support to this recommendation.

H. *Unassignable variation*

In spite of the efforts to age-standardize the mortality data and to remove socio-economic variables through an income adjustment, a considerable amount of variation remains in the data. The mortality data in a sample of 57 counties across 9 States were subjected to a binomial analysis of variance according to methods described by Cochran (1943). This analysis permits a separation of the "error

variance" (between-county variance) into binomial and extraneous components. About 70 percent of the variance was defined as binomial in origin, which explains why data adjustments sometimes seem futile. This finding is also encouraging, however, since it indicates that we are dealing with a variance that is close to the "background noise" level in the population. Standard errors were therefore derived by customary binomial techniques.

This analysis served an additional purpose: it resolved the question of weighting functions for certain analyses and regressions. The 70 percent binomial variance figure permits the use of partial weighting (Cochran, 1943). The array of values was ordered according to sample size and then individual mortality values were weighted by 'n' through the first one-third of the order. The value of 'n' at that sample was then used as a constant weight for the balance of the observations. This proved very successful in preventing overemphasis from values that came from large metropolitan areas such as Denver and Albuquerque. The technique of partial weighting would seem to deserve more attention in epidemiologic analysis.

BIRTH WEIGHT

The published birth weight data are presented at the county level in three categories: "2,500 grams or less," "2,501 grams or more" and "not stated." In the Grahn-Kratchman study, the data were reexpressed in terms of "percent live births at 2,500 grams or less." Because of this rather coarse classification and because of the somewhat secondary application of the birth weight data in the original study, they were not subjected to the same scrutiny as the neonatal mortality data and, therefore, statistical adjustments were not made.

It is difficult to assess the magnitude of any error that might have arisen through this use of essentially "raw" data. At the moment, one can only note that the weight data relationships with altitude (table 2) are entirely consistent with expectations in view of known relations between mortality and weight. A more complete analysis of birth weight variation is presently underway, as will be noted in the "DISCUSSION." For the purposes of this report, the following itemization is given on several sources of variation in weight that should be considered for standardization procedures. Unfortunately, however, not all of these can be controlled through the use of only the published vital statistics and census data.

A. *Race*

Examination of the published data clearly indicates the existence of a significant difference in birth weight between the white and nonwhite classes; the latter are much lighter at birth. This is probably predominantly of environmental rather than genetic origin, though

735-712 O—65——22

there is evidence for the existence of a small amount of genetic variation in birth weight (Penrose, 1954; Morton, 1955).

B. *Maternal age*

Birth weight increases nonlinearly with maternal age, though the difference from 17 to 42 years of age is less than 200 grams. Most of this increase occurs between 20 and 30 years.

C. *Duration of gestation*

Birth weight doubles in the last trimester and increases by several hundred grams between the 38th and 40th weeks. Even minor discrepancies in recording practices can thus lead to errors of 5 percent or more in the weight-gestation length date.

D. *Plurality or multiple births*

The published data are not presented in such a way that this variable can be examined in the smaller political subdivisions. Birth weight declines sharply with increasing plurality, but plural births are probably of a frequency such that they can be ignored without serious consequence when dealing with large population groups.

E. *Sex*

In the case of both neonatal mortality and birth weight data, separation by sex is not provided down to the county level of reporting. This is unfortunate since sex differences in death rate and sex ratio of live births would be of particular value for genetic studies (Schull and Neel, 1958).

F. *Socio-economic factors*

No direct test was made of the existence of possible relationships between birth weight and the socio-economic factors explored with neonatal mortality. However, since the probability of neonatal death is inversely related to weight, and in part certainly a function of weight, the previously discussed factors of income and education are undoubtedly important. These may well define the level of prenatal care that influences fetal growth and therefore early mortality

DISCUSSION

One must now ask—"What is the genetic worth of these data?" Obviously they cannot aid in the determination of genetic mechanisms and their use in genetic studies must be restricted to those endpoints already known to have a genetic etiology. Grahn and Kratchman (1963) have suggested that about 20 percent of the deaths in the neonatal period may be attributed to genetic factors. Most of this mortality is due to congenital malformations and erythroblastosis. Thus, for studies in radiation epidemiology, neonatal mortality could be employed if several conditions were met. These are: uniform

accuracy of diagnosis of cause of death; the existence of genetic equilibrium for the induction and elimination of recessive deterimental genes in populations at higher than average levels of background radiation; and, finally, the assumption that all the excess or difference in mortality in class or classes of death under study is due to genetic factors.

At present, the first condition cannot be met but again recommendations were made at the Ninth Biennial Conference on Records and Statistics to bring about some standardization of reporting of at least the malformation deaths. The second condition probably cannot be met anywhere in the U.S., but certainly not in the western mountains whose populations are comparatively new and of heterogeneous racial origin. The third condition cannot be prejudged but must be ascertained. The ability to detect the induction of phenocopies by some environmental factors other than radiation, but concurrent with it, is a most critical feature of a sound public health approach to radiation epidemiology. The Grahn-Kratchman study is a case in point, since it detected the subtle influence of a common environmental factor—a varying oxygen partial pressure—while seeking effects attributable to terrestrial radiation.

It seems unlikely that the neonatal mortality data can be used to measure the effects of genetic selection and adaptation as the author had suggested in the previous report (Grahn and Kratchman, 1963). Figure 1 indicates that a tremendous improvement in the viability of the newborn has occurred in just 20 years, less than a generation. In other words, the general level of public health exerts a predominant effect. Similarly, Utah, which is an exception to the other Mountain States by having a lower than average infant death rate, has remained uniformly below the U.S. average in the 1940–1960 period. In spite of this, the Grahn-Kratchman study noted regressions of death rate and birth weight on altitude within the State that are consistent with those reported for the other Mountain States. Thus, an environment-specific mortality does exist within Utah. Its effect, if any, on the genetic quality of the population could only be detected by comparison of the regressions on altitude derived from data from different generations. However, this would also involve the comparison of a large number of confounded but extraneous variables.

The comparative analysis of "fertility ratios" and neonatal death rates may also have some genetic significance. This ratio, the number of children under 5 years per 1,000 women between the ages of 15 and 49, is positively correlated with neonatal mortality among the nine census divisions of the U.S. as well as the counties of Colorado. A careful examination of the data over wide extremes of mortality may reveal the extent of the compensatory reproductive capacity of the human female in a modern society, but again the number of

extraneous variables may be too large to warrant meaningful comparisons.

The analysis of birth weights is still incomplete and the published data are not adequate for careful study. At best, they are presented in 500-gram class intervals which is too coarse a measure of the frequency distribution to reveal small changes in the higher moments. Figure 3 shows the distributions derived from the data published in 500 gram intervals. Several features are worthy of mention. Firstly, the whole distribution of birth weights shifts to lower values at higher altitudes. Secondly, there is no clear evidence in these data for or against the existence of an adaptive norm for the human birth weight. The latter has been suggested to exist over the intermediate range of birth weights (Stern, 1960) and one might, therefore, expect a less pronounced shift in the means. The point cannot be completely evaluated since the present data contain the metrical bias depicted in figure 4. The positive correlation ($r=+0.694$; $P<.05$) between mean and variance may be a simple function of the multiplicative quality of growth and weight data and can be easily rectified by a log transformation. The present data do not warrant this effort since the class interval is too broad.

An arrangement has been made with the Department of Health, State of Colorado, to obtain their live birth and fetal and neonatal death records for the period 1959 to 1962. This will involve about 175,000 births and should permit a careful study of the birth weight distribution and the determination of the third and fourth moments with and without the use of a transform. It will then be possible to search more realistically for a negative skewness, for example, that could be evidence for a segregation away from an adaptive norm. The hint of kurtosis with decreasing mean (figure 3) can also be carefully studied. Colorado counties range from 3,500 to over 10,000 feet in mean altitude and this State alone may answer many questions concerning the altitude effect on fetal growth and birth weight.

One other opportunity may also exist to determine the specific effects of altitude as mediated through birth weight. There have been several recent reports on the relation between birth weight and cigarette smoking by the mother during gestation (Herriot *et al.*, 1962). The results suggest that smoking mothers have a higher frequency of infants in the immature class and this has been interpreted as a function of a mild hypoxia due to the higher levels of carboxyhemoglobin in the blood (Oski, 1962). Thus, the tissue effect in the fetus may be the same as living at a higher altitude. How, then, do the mortality statistics compare among smokers vs. nonsmokers in sea level vs. mountain residents? Such a study might permit by inference the isolation of a cosmic or general environmental radiation effect. Also, any selective effect of hypoxia on the immature newborn could be isolated as a specific mountain-residence factor and associated

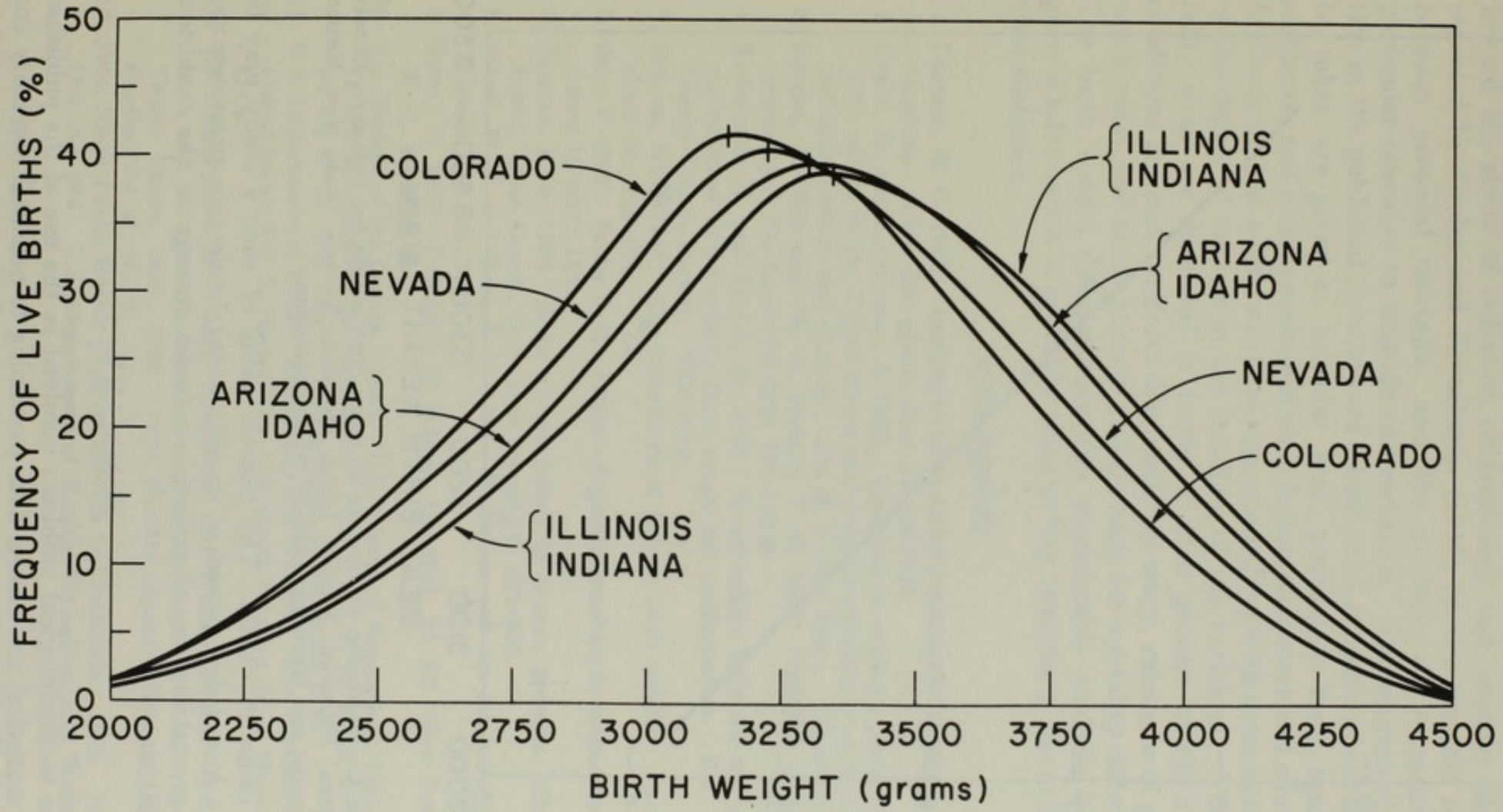


FIGURE 3.—Distribution of live births for indicated States, 1956–57, white population.

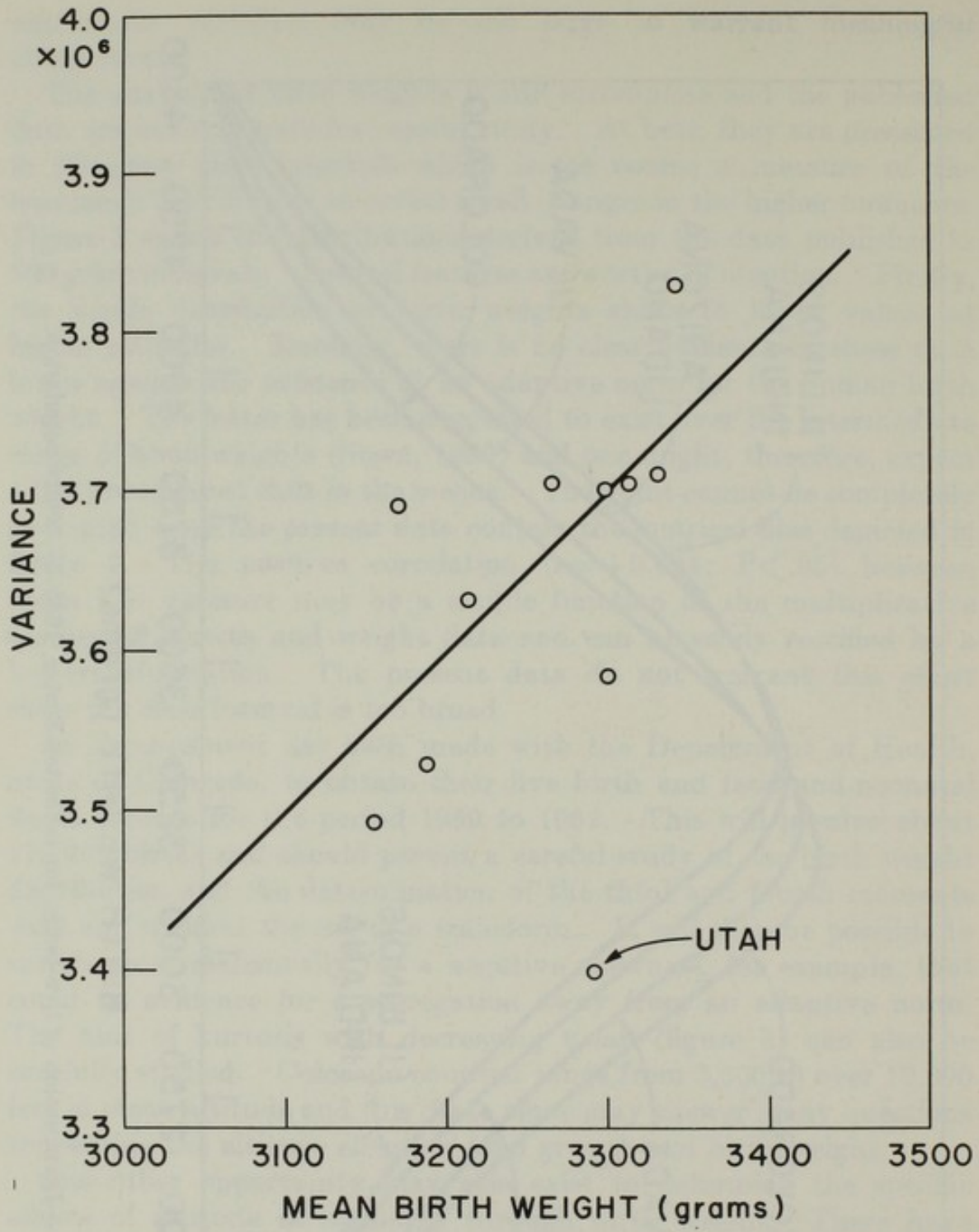


FIGURE 4.—Regression of variance on mean birth weight for Delaware, Illinois-Indiana, Nebraska, Arizona, Idaho, Montana, Nevada, Utah, New Mexico, Colorado, and Wyoming, 1956–57, white population.

with cause of death. The genetic value of such a study may lie in the knowledge gained of specific natural selection pressures and of any contribution of radiation-induced damage in the mountain populations.

SUMMARY

The methodological problems involved in the use of the published vital statistical and census data pertaining to neonatal mortality and

birth weight have been enumerated. The discussion has been directed toward the study of radiation epidemiology and employs the data reported by Grahn and Kratchman (1963) on the possible relation between neonatal mortality and the geologic environment with particular reference to altitude and radiation. The sources of variation in the published data toward which some measure of control can be taken are given as calendar year, population base, age, race, residence, and certain socio-economic factors as income and education. Methods of data adjustment are suggested along with some indication of the degree of success in the handling of the variables. The mortality data are discussed in terms of their genetic worth if three conditions are met: uniform diagnostic accuracy, existence of genetic equilibrium, and assumption of genetic basis for mortality differential. The birth weight data still require considerable analysis but the genetic information is probably small in any analysis of the published vital statistics.

REFERENCES

1. Cochran, W. G. 1943. Analysis of variance for percentages based on unequal numbers. *Amer. Stat. Assoc. Jour.* 38: 287–301.
2. Grahn, D. and Kratchman, J. 1963. Variation in neonatal death rate and birth weight in the United States and possible relations to environmental radiation, geology and altitude. *Am. J. Human Genet.* 15: 329–352.
3. Herriot, A., Billewicz, W. Z., Hytten, F. E. 1962. Cigarette smoking in pregnancy. *The Lancet* 1(7233): 771–773.
4. Kratchman, J. and Grahn, D. 1959. Relationships between the geologic environment and mortality from congenital malformation. U.S. Atomic Energy Comm. Report. TID–8204.
5. Morton, N. E. 1955. The inheritance of human birth weight. *Ann. Hum. Genet.* 20: 125–134.
6. Oski, F. 1962. Letter to the Editor—Cigarette smoking in pregnancy. *The Lancet* 1(7240): 1188.
7. Penrose, L. S. 1954. Some recent trends in human genetics. Proc. IX Intern. Cong. Genet., Caryologia (Suppl.) 6: 521–530.
8. Schull, W. J. and Neel, J. V. 1958. Radiation and the sex ratio in man. *Science* 128: 343–348.
9. Stern, C. 1960. Principles of Human Genetics, 2d Ed. San Francisco: W. H. Freeman Co.
10. U.S. Bureau of the Census. 1953. U.S. Census of Population: 1950. Vol. II, Characteristics of the population. Parts 1, 3, 5, 6, 16, 27, 28, 31, 34, 41, 43, 44, 47, 50, U.S. Gov't. Printing Office, Washington, D.C.
11. U.S. Department of Health, Education, and Welfare. Vital Statistics of the United States, Vol. I., 1950–1957. U.S. Gov't. Printing Office, Washington, D.C.
12. U.S. Department of Health, Education, and Welfare. 1958. Weight at birth and survival of newborn by age of mother and total birth order: United States, early 1950. Vital Statistics-Special Reports, Selected Studies, Vol. 47, No. 2.
13. Proc. Ninth Public Health Conf. on Records and Statistics. 1962. PHCRS Doc. No. 574. U.S. Department of Health, Education, and Welfare.

PART IV—Concluded

Malignant Diseases

Familial Factors in Lung Cancer and Smoking

By George K. Tokuhata, Ph. D., Dr. P.H.,
Division of Chronic Diseases, Public Health Service, U.S. Department of Health, Education, and Welfare[1]

A number of investigators have postulated that genetic factors may play an important role in the etiology of lung cancer (Norton, 1959; Goodhart, 1959; Lacassagne, 1956; Shettles, 1958; Berkson, 1955). Likewise it has been suggested that there may be a biological basis for the habit of smoking (Heath, 1958; Parnell, 1951; Lilienfeld, 1959). Several laboratory studies (Cohen and Thomas, 1962; Thomas and Cohen, 1960) have reported that the smoking habit may be related to the ABO blood type and the ability to taste P.T.C. which are known to be genetically determined. A higher degree of concordance with respect to smoking habits was found among monozygotes than among dizygotes in a few twin studies (Fisher, 1958; Friberg, 1959). An association between constitutional typology and the smoking habit has also been demonstrated (Seltzer, 1962). In addition, several epidemiologic studies (Bothwell, 1959; Horn *et al.*, 1959; Salber and MacMahon, 1961) have shown that smokers tend to aggregate in families.

Since lung cancer and cigarette smoking are closely related to each other, one can speculate that there may be some unidentified factors at the cellular level which may be responsible for the incidence of lung cancer and the distribution of smoking habits in our population. Fisher (1958) hypothesized that both lung cancer and smoking habits are determined by a "common genotype"; this implies no direct causal relationship between the two.

OBJECTIVES OF THE STUDY

The epidemiology of lung cancer and of cigarette smoking in the same family population has not been fully explored. This is a synthe-

[1] Present address: St. Jude Research Hospital, Memphis, Tennessee.

sized report of two separate papers, one dealing with the *familial aggregation of lung cancer* (Tokuhata and Lilienfeld, 1963) and the other dealing with the *familial aggregation of smoking* (Tokuhata, 1963). In the present report we will particularly emphasize the methodological aspects of the study. In an attempt to elucidate the possible importance of the heredofamilial factor, the data will be analyzed in such a way as to equalize some of the known influences of the host and environmental characteristics which are associated with lung cancer and/or smoking habits.

The specific objectives of the present paper are threefold: First, to illustrate a method by which familial aggregation of lung cancer may be demonstrated while adjusting for the influences of cigarette smoking and other factors. Second, to illustrate a method by which familial aggregation of cigarette smoking may be demonstrated while adjusting for the effects of the familial history of lung cancer and other factors. And, third, to attempt to develop some tentative inferences with respect to different biological predispositions to the cigarette smoking habits and to cancer of the lung.

METHODS

I. *Selection of Lung Cancer Probands and Index Controls:*

Lung cancer patients used as *probands* (*index cases*) in this study had histologically confirmed diagnoses and were reported from nine major hospitals in Baltimore during the 1960–61 period. No distinction was made with respect to the specific tumor type. The study was limited to Caucasians. These probands, including both in- and out-patients as well as private and ward patients, were estimated to represent at least 90 percent of all cases of lung cancer prevalent in Baltimore City and Baltimore County at the time of the study. For purposes of comparison, a group of apparently healthy people were selected as *index controls* who were matched for race, sex, age (within a range of plus or minus 5 years), and usual residence. This was done by systematically screening adjacent households in the same block as the proband until a living person was located who had the necessary matching criteria.

Of 317 cases of lung cancer reported during the study period, 47 were eliminated from the study because the necessary information on relatives could not be obtained and/or most of the patients' relatives had died in non-English-speaking countries. Thus, for the remaining 270 probands an equal number of "neighborhood controls" were selected. Having interviewed these controls, 27 were found to be unusable for the same reason; these were then replaced by those for whom adequate information was ascertained. The method of

obtaining such replacements was the same as that used in the initial selection of controls. The numerical and percentage distributions by age groups and sex of the index subjects thus selected are given in table 1. As shown in this table, the probands and index controls are comparable with respect to age and sex composition.

TABLE 1.—*Distribution of Case and Control Index Subjects by Sex and Age Group*

Age group	Number of males		Number of females		Percent distribution of males		Percent distribution of females	
	Case	Control	Case	Control	Case	Control	Case	Control
Under 40	1	3	0	1	0.4	1.2	0.0	5.0
40–44	5	14	1	1	2.0	5.6	5.0	5.0
45–49	18	14	2	3	7.2	5.6	10.0	15.0
50–54	37	34	2	1	14.8	13.6	10.0	5.0
55–59	58	50	2	2	23.2	20.0	10.0	10.0
60–64	53	42	5	1	21.2	16.8	25.0	5.0
65–69	40	43	2	4	16.0	17.2	10.0	20.0
70–74	30	27	4	1	12.0	10.8	20.0	5.0
75–79	6	19	1	3	2.4	7.6	5.0	15.0
80 and over	2	4	1	3	0.8	1.6	5.0	15.0
All ages	250	250	20	20	100.0	100.0	100.0	100.0

II. *Collection of Data:*

A. *Interview Questionnaire:*

The necessary preliminary data were obtained by interviewing the index subjects and their immediate family members. The questionnaire contained the following items: (1) lifetime smoking experience including the age at which regular (daily) smoking was started, and the form, amount and duration of smoking, and date of birth of the index subject; (2) lifetime smoking experience including the form and regularity of smoking, date of birth, sex, living status, and current address of each of the children, siblings, parents, and spouses; and (3) name of each of the deceased relatives including information regarding place, year, and age at death. No individuals under 20 years of age were considered in this study.

B. *Smoking History:*

To ascertain or validate the smoking information of the relatives of the index subjects, a post card was used which contained simple questions regarding lifetime history of smoking, particularly with respect to the form of smoking adopted such as cigarettes, cigars, pipes, or others, and the regularity of smoking such as "occasionally" or "daily". A series of possible answers, all mutually exclusive, in each smoking category, were provided on each card for a respondent to check. An effort was made to obtain the name and address of at least one informant for each of the deceased relatives whose usual address was not the same as the index subject being interviewed or for each of those living relatives whose addresses were not available.

The lung cancer probands had 2,018 living and deceased relatives whose smoking information was to be ascertained. To each of the living relatives including prospective informants the smoking questionnaire card was mailed with an accompanying letter. Of the total smoking questionnaire cards mailed, 1,392 (60 percent) were returned with the necessary information, 69 (3 percent) were returned without information, and the remaining 557 (28 percent) were not returned. To those who did not return the cards, a series of followup cards were mailed. Frequent telephone contacts were also made for the same purpose. To evaluate the validity of smoking information the following criteria were used: 1) the information obtained by direct contact was always used; 2) if the information given by an informant did not agree with the information originally obtained from the interviewed index subject, another informant was used. In this situation, the information supported by two sources was finally used. Through these means the necessary smoking data were ascertained for an additional 424 (21 percent) individuals. Thus, information on smoking was complete for a total of 90 percent of the 2,018 relatives of the probands.

Likewise, 1,903 smoking questionnaire cards with accompanying letters were mailed to the relatives of the 270 controls. Of these cards, 1,449 (76 percent) were returned with information, 16 (1 percent) were returned without information, and the remaining 438 (23 percent) were not returned. However, by the same followup methods the necessary information was obtained for an additional 249 (13 percent) individuals. Thus smoking data were complete for 89 percent of the 1,903 relatives of the controls.

C. *Cause of Death:*

An attempt was made to locate a death certificate for each deceased family member to ascertain the cause(s) of death. Of 1,676 deaths, including spouses, reported in this study, 1,543 (92 percent) death certificates were located. Of the remaining 133 deaths, for which death certificates were not located 14 (0.8 percent) had occurred during a preregistration period in different States, while 119 (7.1 percent) had insufficient information regarding the time, place, and/or name of the deceased.

With respect to the geographic distribution of the deaths, the majority (73.7 percent) had occurred in the State of Maryland, mostly within the city of Baltimore, and the number decreased sharply as the distance from Maryland increased, e.g., 122 (7.3 percent) in Pennsylvania, 57 (3.4 percent) in Virginia, 45 (2.7 percent) in West Virginia, etc. There were also 19 (1.1 percent) deaths that occurred in non-English-speaking countries for which we did not attempt to obtain death certificates.

III. *Age Adjustment on Mortality Among Relatives:*

To adjust for the effect of age variation on mortality the *observed mortality experience* of the proband relatives was compared with the *expected mortality experience* of the same relatives as estimated on the basis of the age-specific mortality ratios of the control relatives. To compute the age-specific mortality ratio in a given category of relatives it is necessary to determine the average number of such relatives exposed to risk in each age group. The formula used for this calculation is as follows:

Ni (average number of relatives in a given category exposed to risk in *i*th age group)=Oxi−½ Qxi.

Oxi=Number of all relatives in a given category who have reached the youngest age in *i*th group.

Qxi=Number of relatives, both living and dead, in *i*th age group.

The computation of the expected number of deaths among proband relatives may be symbolically represented as follows:

Expected number of deaths among proband relatives

$$=\sum_{1}^{n}\frac{Dci}{Nci}Npi.$$

Dci=Number of deaths in *i*th age group of the control relatives.

Nci=Average number of control relatives exposed to risk in *i*th age group.

Npi=Average number of proband relatives exposed to risk in *i*th age group.

n=Number of age groups.

i=A given age group (e.g., 30–34, 35–39, etc.).

As an illustration of the above method of computation the data for *sisters* using all-cause mortality are shown in table 2. The same method of analysis is applicable to any other category of relatives as well as any specific cause of death.

TABLE 2.—*Method Used To Calculate the Expected Number of Deaths from All Causes Among Sisters by Application of the Control Experience*

Age at report	Control				Case			
	Number of relatives	Average number exposed to risk	Total number of deaths	Proportion of deaths	Number of relatives	Average number exposed to risk	Total number of deaths	Expected number of deaths
30–39	35	446	13	0.0291	48	537	23	15.63
40–49	99	379	16	0.0422	91	468	16	19.75
50–59	137	261	22	0.0843	186	329	42	27.73
60–69	120	132	18	0.1364	164	154	24	21.01
70–79	63	41	14	0.3415	61	42	12	14.34
80+	9	5	2	0.4000	11	6	5	2.40
Total	463		85		561		122	100.86

NOTE: Age under 30 or unknown is excluded.

I. *Numerical Composition of Relatives:*

Before presenting the results of the analysis of the lung cancer mortality and the smoking habits among relatives, a brief comparison of the number of relatives of the probands with that of the controls is in order. The numerical and percentage distribution of mothers, fathers, sisters, and brothers *by year of birth* are given in table 3. It is clear from this table that parents and siblings of the probands were quite comparable regarding year of birth with the counterparts of the controls. The distribution of index subjects *by number of siblings and children* is presented in table 4. As shown in this table, there were no significant differences in sibship and in family size between the two groups of index subjects being compared. Although the number of probands and index controls was not very large, statistical analyses of the mortality and smoking experience among the relatives were based on approximately 50,000 average "person-years" in each of the proband and control family groups.

TABLE 3.—*Numerical and Percentage Distribution of Relatives by Year of Birth*

Year of birth	(A) Distribution of parents							
	Numerical distribution				Percentage distribution			
	Mother		Father		Mother		Father	
	Case	Control	Case	Control	Case	Control	Case	Control
Before 1850	8	10	12	22	3.0	3.7	4.4	8.1
1850–1859	26	29	47	41	9.6	10.7	17.4	15.2
1860–1869	86	68	95	83	31.9	25.2	35.2	30.7
1870–1879	90	81	74	63	33.3	30.0	27.4	23.3
1880–1889	46	52	33	44	17.0	19.3	12.2	16.3
1890–1899	10	25	3	10	3.7	9.3	1.1	3.7
1900 and after	3	4	2	3	1.1	1.5	0.7	1.1
Unknown	1	1	4	4	0.4	0.4	1.5	1.5
Total	270	270	270	270	100.0	100.1	99.9	99.9

Year of birth	(B) Distribution of siblings							
	Numerical distribution				Percentage distribution			
	Sister		Brother		Sister		Brother	
	Case	Control	Case	Control	Case	Control	Case	Control
Before 1870	4	0	4	3	0.7	0.0	0.7	0.6
1870–1879	14	22	21	20	2.4	4.6	3.7	3.8
1880–1889	83	59	92	75	14.1	12.3	16.3	14.4
1890–1899	176	124	164	140	29.9	25.8	29.1	26.8
1900–1909	171	136	156	129	29.1	28.3	27.7	24.7
1910–1919	100	93	89	106	17.0	19.4	15.8	20.3
1920 and after	38	37	37	39	6.5	7.7	6.6	7.5
Unknown	2	9	1	10	0.3	1.9	0.2	1.9
Total	588	480	564	522	100.0	100.0	100.1	100.0

TABLE 4.—*Distribution of Case and Control Subjects by Number of Siblings and Children*

Number of sibs* (exclusive of subject)	Number of families		Percent distribution		Number of children*	Number of families		Percent distribution	
	Case	Control	Case	Control		Case	Control	Case	Control
0	19	23	7.0	8.5	0	56	43	22.4	17.0
1	30	35	11.1	13.0	1	60	51	24.0	20.1
2	29	40	10.7	14.8	2	37	69	14.8	26.6
3	33	48	12.2	17.8	3	46	42	18.4	16.6
4	36	34	13.3	12.6	4	22	24	8.8	9.3
5	35	31	13.0	11.5	5	16	15	6.4	5.8
6	37	18	13.7	6.7	6	4	7	1.6	2.7
7	19	17	7.0	6.3	7	4	6	1.6	2.3
8+	32	24	11.9	8.9	8+	5	2	2.0	0.8
Total families	270	270	99.9	100.1	Total families	250	259	100.0	101.2
Total sibs	1,152	1,002			Total children	533	587		
Average sibs	4.3	3.7			Average children	2.1	2.3		

*Those under 20 years of age are excluded.

Significance Test: Sibs $X^2=13.4$ (d.f.=8) not significant.
Children $X^2=14.9$ (d.f.=8) not significant.

II. *Familial Aggregation of Lung Cancer:*

In analyzing lung cancer as cause of death, both categories, *i.e.*, cancer of the lung and bronchus specified as primary and that not specified as primary or secondary were included. However, those specified as secondary or metastatic were excluded. In this analysis, those with unknown ages or under 30 years of age and those whose smoking information was not available were excluded.

To determine the presence of familial aggregation of lung cancer, the lung cancer mortality among relatives of a group of lung cancer probands was compared with that of relatives of a comparable group of individuals selected as *index controls*. In this comparison it is necessary to take into account simultaneously the effects of several host and environmental factors such as cigarette smoking, sex, age, and generation (category of relative). This was done by applying the *method of age adjustment*, as described earlier, to various groups of relatives cross-classified according to sex, generation, and smoking status. Specifically, the observed number and proportion of lung cancer deaths in each of the four categories of proband relatives (mothers, fathers, sisters, and brothers) who smoked and who did not smoke cigarettes was compared with the number and proportion of such deaths that would be expected among the same relatives if the sex-age-generation-smoking specific mortality ratios of the control relatives had prevailed. The results of such analyses are summarized in table 5.

As shown in this table, the observed value was greater than expected in all categories of the proband relatives, including both males and females as well as both smokers and nonsmokers. The overall difference for all relatives being considered was statistically significant, indicating that the proband relatives had a significantly increased risk

735-712 O—65——23

TABLE 5.—*Observed and Expected Numbers, and Percentages of Lung Cancer Deaths among Smokers and Nonsmokers by Specific Category of Relatives*

Case relatives	Nonsmokers		Smokers		Smokers and nonsmokers	
	Observed number of Lung Cancer	Expected number of Lung Cancer	Observed number of Lung Cancer	Expected number of Lung Cancer	Observed number of Lung Cancer	Expected number of Lung Cancer
Mothers	5	2.0	0	0.0	5	2.0
Sisters	3	1.2	0	0.0	3	1.2
Total females	8	3.2	0	0.0	8	3.2
Fathers	3	0.0	2	1.1	5	1.1
Brothers	1	0.0	21	9.0	22	9.0
Total males	4	0.0	23	10.1	27	10.1
Total	12	3.2	23	10.1	35	13.3

Case relatives	Nonsmokers			Smokers			Smokers and nonsmokers		
	Number relatives	Percent Lung Cancer		Number relatives	Percent Lung Cancer		Number relatives	Percent Lung Cancer	
		Observed	Expected		Observed	Expected		Observed	Expected
Mothers	259	1.9	0.8	7	0.0	0.0	266	1.9	0.8
Sisters	405	0.7	0.2	155	0.0	0.0	560	0.5	0.2
Total females	664	1.2	0.5	162	0.0	0.0	826	1.0	0.4
Fathers	190	1.6	0.0	70	2.9	1.4	260	1.9	0.4
Brothers	188	0.5	0.0	350	6.0	2.6	538	4.1	1.7
Total males	378	1.1	0.0	420	5.5	2.4	798	3.4	1.3
Total	1,042	1.2	0.3	582	4.0	1.7	1,624	2.2	0.8

	Significance tests		
	Nonsmokers	Smokers	Smokers and nonsmokers
Both sexes	P=0.004	P=0.01	P=0.0006
Males	P=0.02	P=0.01	P=0.003
Females	P>0.05	Undetermined	P>0.05

of dying from lung cancer either in association with, or independently of, the history of cigarette smoking. As an indication of the possible influence of the environmental factor common to members of married pairs, spouses of the probands and those of the index controls were compared with respect to lung cancer mortality. There were no significant differences in such mortality among spouses, including both wives and husbands.

Since the *index controls* were selected within the adjacent neighborhood and the majority of the relatives were also residing in the same geographic areas, the possible effects on lung cancer of such additional environmental factors as atmospheric pollution, socioeconomic status and ethnic background are assumed to be fairly

homogeneous. From the totality of these observations it would appear that some inherent factor may play an important role in the etiology of lung cancer.

III. *Analysis of the Data on Smoking:*

It has been established that the lung cancer patients are much more likely to be cigarette smokers than those who do not have such disease. According to our data, 98.4 percent of the male probands were actually cigarette smokers. This is significantly greater than 81.6 percent that would be expected if the age-specific proportions of cigarette smokers among the male index controls were applied. From this observation several questions may be asked regarding the smoking habits of relatives of the index subjects. First, are the relatives of the probands also more likely to be smokers than the counterparts of the controls? Second, is the smoking pattern in the families of the lung cancer group different from that of the control group? And, third, is there an association between the smoking habits of offspring and that of parents? We will attempt to answer these questions with particular emphasis on the application of several different methods that have been developed in epidemiology and genetics.

It should be noted that the data on smoking among relatives were not ascertained through smokers. Although the majority of the lung cancer probands were smokers, the difference in the proportion of smokers and nonsmokers among the index controls was much less marked. In computing the proportion of smokers, both living and deceased individuals were considered; however, those for whom smoking information and/or age was not given were excluded.

A. *Inter-Familial Aggregation of Smokers:*

To determine the presence of *inter-familial aggregation* of smokers, i.e., excess frequency of smokers among the proband relatives as compared with the control relatives, it is necessary to take into account the possible differences in sex, age, and generation of the two family populations which may be associated with the smoking habits. For this purpose we computed the number of smokers that would be expected among the proband relatives by applying the age-sex-generation specific proportion of smokers among the control relatives. The expected number of smokers was then compared with the observed number within the proband relatives. A summary of these calculations is given in table 6.

As shown in this table, the observed value was significantly greater than expected in both sexes, *i.e.*, both male and female relatives of the probands manifested a greater risk of being smokers than those of the controls. The data further show that this excess frequency of smokers was actually accounted for by siblings of the index subjects.

TABLE 6.—*Observed and Expected Number of Cigarette Smokers Among Case Relatives*

	Number of relatives	Number of smokers		Significance test
		Observed	Expected*	
Fathers	261	70	58.4	P>0.05
Brothers	560	363	322.5	P=0.008
Sons	215	141	150.8	P>0.05
All males	1,036	574	531.7	P=0.05
Mothers	268	7	5.4	P>0.05
Sisters	585	160	129.9	P=0.04
Daughters	199	116	91.1	P=0.01
All females	1,052	283	226.4	P=0.005
Parents	529	77	63.8	P>0.05
Siblings	1,145	523	452.4	P=0.02
Children	414	257	241.9	P>0.05
All relatives	2,088	857	758.1	P=0.0009

*Computed on the basis of the age-specific smoking frequency of the control relatives.

NOTE: Those with unknown smoking status, unknown ages, or those under 20 years of age are excluded.

A similar trend was also observed among parents and children; however, the difference was not statistically significant.

B. *Intra-Familial Aggregation of Smokers:*

The intra-familial aggregation of smokers was determined by comparing the proportion of smokers among those relatives whose index subjects are also smokers with that among those relatives whose index subjects are nonsmokers. This comparison was possible in each of the proband and control family populations as well as in the combined population.

The analysis of intra-familial aggregation of smokers was made for parents, siblings, and children of the index subjects and the results are summarized in table 7. Let us first consider the combined population. The data clearly indicate that there is significant aggregation of smokers in families. Specifically, those parents and siblings whose index subjects are smokers were much more likely to be smokers than those whose index subjects are nonsmokers. Although more than twice as large a proportion of the male as the female relatives are smokers, the same mode of aggregation was revealed in both sexes.

The data also show a striking difference when the proband and control families were considered separately. As mentioned earlier, the overall frequency of smokers among the proband relatives was significantly greater than that among the control relatives. No such relationship was found among spouses of the index subjects. It is of interest to note that about 40 percent of proband relatives reported a history of regular smoking regardless of the smoking status of the index subjects. This generally increased and "diffuse" frequency of smokers among the proband relatives suggests that there may be

TABLE 7.—*Percentage of Cigarette Smokers among Relatives, According to Cigarette Smoking Status and Case Control Identity of the Index Subject*

Category of relatives	Case group		Control group		Combined group	
	Index subjects who are smokers	Index subjects who are nonsmokers	Index subjects who are smokers	Index subjects who are nonsmokers	Index subjects who are smokers	Index subjects who are nonsmokers
Father	27.9	15.0	30.4	12.0	28.9	12.5
Mother	2.8	0.0	3.1	0.0	2.9	0.0
Brother	65.0	62.9	66.9	46.7	65.7	49.0
Sister	27.6	20.7	27.4	14.1	27.5	15.0
Son	66.0	63.0	72.8	67.6	68.7	66.7
Daughter	58.3	58.3	47.1	42.3	53.7	45.3
Total males	55.7	51.2	58.3	43.1	56.7	44.4
Total females	26.8	27.4	24.7	17.5	26.0	19.1
Total parents	15.2	7.5	16.7	6.0	15.8	6.2
Total siblings	45.8	43.8	47.6	31.5	46.4	33.2
Total children	62.3	60.8	60.2	55.0	61.4	56.2
Total	41.1	40.0	41.8	30.7	41.3	32.2

NOTE: Percentage of smokers is computed against those for whom smoking information is given.

Significance test

Case group		Control group		Combined group	
Total relatives	$P>0.05$	Total relatives	$P<0.0001$	Total relatives	$P<0.0001$
Males	$P>0.05$	Males	$P<0.0001$	Males	$P<0.0001$
Females	$P>0.05$	Females	$P=0.005$	Females	$P=0.002$
Parents	$P>0.05$	Parents	$P=0.0002$	Parents	$P<0.0001$
Siblings	$P>0.05$	Siblings	$P<0.0001$	Siblings	$P<0.0001$
Children	$P>0.05$	Children	$P>0.05$	Children	$P>0.05$

something inherent in the biological makeup common to all "blood" relatives of the probands.

It has been pointed out that the control relatives were much less likely to be smokers than the proband relatives as a whole. The data further show that approximately 40 percent of those control relatives whose index subjects were smokers, as compared with about 30 percent of those control relatives whose index subjects were nonsmokers were actually smokers; a difference which is statistically significant. This "selective" intra-familial aggregation of smokers may be due to genetic factors, or common environmental factors, or both.

C. *Parental Smoking and Smoking among Offspring:*

In analyzing the smoking behavior of offspring in relation to that of parents, two generations, i.e., parents and siblings of the index subjects, were considered. Since a number of children of the index subjects were still under the age of 20, this generation was not included.

The number of male and female offspring and the proportion of smokers according to the smoking status of their parents are given in table 8. The relative frequency of matings in which both parents are smokers was about the same in the proband and control families. However, the like frequency in which neither parent is a smoker was much more common in the control than in the proband families.

TABLE 8.—*Number and Proportion (%) of Cigarette Smokers among Offspring According to Parental Smoking Habits*

(A) CASE FAMILIES

Parental mating type	Number families	Male offspring			Female offspring			All offspring		
		Total	Smokers	Percent smokers	Total	Smokers	Percent smokers	Total	Smokers	Percent smokers
Both smokers	7	16	15	93.8	14	11	78.6	30	26	86.7
One smoker	171	509	380	74.7	412	119	28.9	921	499	54.2
Neither smoker	85	274	196	71.5	171	42	24.6	445	238	53.5
Total	*263	799	591	74.0	597	172	28.8	1,396	763	54.7
X^2 Test (d.f.=2)		$X^2=4.28$ $P>0.05$			$X^2=18.67$ $P<0.001$			$X^2=12.73$ $0.001<P<0.01$		

(B) CONTROL FAMILIES

Both smokers	5	14	13	92.9	11	6	54.5	25	19	76.0
One smoker	152	413	258	62.5	268	73	27.2	681	331	48.6
Neither smoker	110	333	184	55.3	213	28	13.1	546	212	38.8
Total	**267	760	455	59.9	492	107	21.7	1,252	562	44.9
X^2 Test (d.f.=2)		$X^2=10.37$ $0.001<P<0.01$			$X^2=20.89$ $P<0.001$			$X^2=21.75$ $P<0.001$		

(C) CASE AND CONTROL FAMILIES

Both smokers	12	30	28	93.3	25	17	68.0	55	45	81.8
One smoker	323	922	638	69.2	680	192	28.2	1,602	830	51.8
Neither smoker	195	607	380	62.6	384	70	18.2	991	450	45.4
Total	530	1,559	1,046	67.1	1,089	279	25.6	2,648	1,325	50.0
X^2 Test (d.f.=2)		$X^2=16.81$ $P<0.001$			$X^2=37.07$ $P<0.001$			$X^2=32.81$ $P<0.001$		

*Seven families are excluded.
**Three families are excluded.

NOTES: 1. Offspring includes index subjects and their siblings.
2. Offspring includes all those 20 years of age or over.
3. Occasional smokers are considered as nonsmokers.
4. Those siblings whose smoking information or whose parental smoking information is not available are excluded.

A statistical analysis indicates that there is a high degree of association between the parental smoking habits and the smoking habits of their offspring. Specifically, in both the proband and control families, children are much more likely to be smokers when parents are in fact smokers. It should be mentioned, however, that from the statistical association such as shown here it is not possible to differentiate the source of influence between the genetic factors and the familial environmental factors.

SUMMARY AND DISCUSSION

In the light of known association between lung cancer and smoking, suspected biological predisposition to lung cancer, and some suggestive evidence supporting the role of a genetic factor in smoking, we have

compared the data on the lung cancer mortality and cigarette smoking habits among parents, siblings, and children of a group of lung cancer probands with those of the counterparts of a group of index controls. This paper has emphasized the methodology as employed in the epidemiologic analysis of family data.

According to the present study, the relatives of the lung cancer probands had an excess risk of dying from the same disease *with or without* the presence of cigarette smoking, a most significant single factor associated with the disease. No such risk was found among spouses of the probands. Further, the risk of developing lung cancer was the same for both men and women who do not smoke cigarettes. In addition to cigarette smoking, the influences of sex, age, and generation on lung cancer were also taken into consideration. Further, since neighborhood controls were used in this study, the possible effects of atmospheric pollution and other common environmental factors are assumed to be relatively equalized. These observations were interpreted as suggesting that certain biological predisposition or genotype may play a role in the pathogenesis of lung cancer.

Our data have also shown that the relatives of the probands are indiscriminatively much more likely to be cigarette smokers than those of the controls. No such difference was indicated among spouses of the index subjects. Further, the intra-familial aggregation of smoking has been demonstrated in the control families. A detailed analysis of the data on smoking suggests that the control families may be composed of those with "high" and those with "low" risk with respect to being smokers whereas the proband families cannot be so classified with almost all families manifesting the same "high" risk.

Although there is a close correlation between lung cancer and cigarette smoking in the same individuals, familial aggregation of lung cancer has been demonstrated independently of smoking, and likewise familial aggregation of smoking has been demonstrated independently of lung cancer. Our data also indicate that the degree of association between lung cancer and the familial factor (history of lung cancer in family) is nearly as strong as that between lung cancer and smoking. From these observations it would appear that even though both smoking habit and lung cancer exist in some individuals, these two entities may not share the same biological basis. Thus, the results of our analysis do not seem to support Fisher's common genotype hypothesis.

There is a question of sex difference in smoking and lung cancer. One possible approach to this question regarding smoking would be in terms of "sex influence." That is, the fact that smoking is more prevalent among men than among women may be in some degree due to socially prescribed norms. It is noteworthy that in recent years such a difference appears to be diminishing. Regarding the sex difference in lung cancer, it has been found that both men and women

are equally liable to this disease if they are nonsmokers, but men are more liable than women if they are smokers. No simple explanation is available for the difference observed among smokers. It is possible that such a difference may be due to unidentified differences in the physiological or biochemical makeup of men and women.

In any epidemiologic study designed to assess the role of genetic factors it is difficult to identify and differentiate the nature and degree of the potential environmental influences common to all family members. An ideal experimental study to investigate the problem of multiple causality in smoking and lung cancer may be designed with a large number of pairs of twins ascertained at birth and followed for as long as 50 years or longer with complete information on each individual regarding major environmental factors associated with each of the two entities. Obviously such a study has far from practical feasibility.

The available data to support the implicated role of a genetic factor in lung cancer and smoking are still incomplete. In a larger study of the familial aggregation of lung cancer it would be possible and desirable to analyze separately adenocarcinoma and other types of bronchogenic carcinoma. This would be particularly important since adenocarcinoma is not associated with smoking. Our present knowledge regarding the behavioral aspects of smoking habits is extremely limited. Future studies specifically designed to clarify the nature and degree of the familial and other environmental influences on smoking will provide an essential empirical basis for better understanding of the problem.

REFERENCES

1. Berkson, J. 1955. The statistical study of association between smoking and lung cancer. *Proc. Mayo Clin.* 30: 319–348.
2. Bothwell, P. W. 1959. The epidemiology of cigarette smoking in rural school children. *Med. Officer.* 102: 125–132.
3. Cohen, B. H., and Thomas, C. B. 1962. Comparison of smokers and nonsmokers. II. The distribution of ABO and Rh(D) blood groups. *Bulletin of the Johns Hopkins Hospital.* Vol. 110, No. 1.
4. Fisher, R. A. 1958. Lung cancer and cigarettes? *Nature* 182: 108.
5. Friberg, L. 1959. Smoking habits of monozygotic and dizygotic twins. *Brit. Med. J.* 1: 1090–1092.
6. Goodhart, C. B. 1959. Cancer-proneness and lung cancer. *Practitioner* 182: 578–584.
7. Heath, C. W. 1958. Differences between smokers and nonsmokers. *AMA Arch. Intern. Med.* 101: 377–388.
8. Horn, D., Courts, F. A., Taylor, R. M., and Solomon, E. S. 1959. Cigarette smoking among high school students. *Amer. J. Public Health.* 49: 1497–1511.
9. Lacassagne, A. 1956. Present-day aspects of research on the etiology of pulmonary tumours; Osler oration. *Canad. Med. Assoc. J.* 75: 894–897.

10. Lilienfeld, A. M. 1959. Emotional and other selected characteristics of cigarette smokers and nonsmokers as related to epidemiological studies of lung cancer and other diseases. *J. Nat. Cancer Inst.* 22: 259–282.
11. Norton, H. W. 1959. Correspondence. *Amer. Scientist* 47: 141–142.
12. Parnell, R. W. 1951. Smoking and cancer. *Lancet* 1: 963.
13. Salber, E. J., and MacMahon, B. 1961. Cigarette smoking among high school students related to social class and parental smoking habits. *Amer. J. Public Health* 51: 1780–1789.
14. Seltzer, C. C. 1962. Why people smoke. *Appl. Ther.* 4: 1023–1024.
15. Shettles, L. B. 1958. Biological sex differences with special reference to disease, resistance, and longevity. *J. Obstet. Gynec. Brit. Emp.* 65: 288–295.
16. Thomas, C. B., and Cohen, B. H. 1960. Comparison of smokers and nonsmokers. I. A preliminary report on the ability to taste Phenylthiourea (P.T.C.) *Bulletin of the Johns Hopkins Hospital,* Vol. 106, No. 4.
17. Tokuhata, G. K., and Lilienfeld, A. M. 1963. Familial aggregation of lung cancer in humans. *J. Nat. Cancer Inst.* 30: 289–312.
18. Tokuhata, G. K. 1963. Smoking habits in lung cancer proband families and comparable control families. *J. Nat. Cancer Inst.* 31: 1153–1171

Do Malignancies Result From Diagnostic and Therapeutic Radiation?

By T. Sterling, Ph. D., E. Saenger, M.D., and R. A. Seltzer,* *College of Medicine, University of Cincinnati, Cincinnati, Ohio*

A number of investigators during recent years have come to the conclusion that radiation in any dose constitutes a hazard (MacMahon, 1962; Simpson *et al.*, 1955, 1957; Stewart *et al.*, 1956, 1958; Stewart, 1961). Specifically it has been claimed that diagnostic and therapeutic doses have caused an increase in certain types of cancer, specifically of leukemia. Proof for these assertions were based on investigations in which classical epidemiological methods, suitable to the study of infections and nutritional diseases, were applied to problems of great complexity for which they may not have been suitable. The epidemiological problems raised by the wholesale importation of traditional epidemiological techniques to the study of so-called chronic diseases have not been appreciated. Traditional statistical and epidemiological procedures worked well for acute infectious diseases, where time and host susceptibility did not play a large role, but when such procedures were used to identify a particular agent as a disease antecedent, there followed a clash of opinions rather than well documented conclusions. We have now reached a point for reevaluation of the present armamentarium and for forging new epidemiological tools. With this new look the term "chronic disease epidemiology" should take on new meaning as a set of complex problems are investigated which require varied and pioneering solutions.

Within this framework, the problem of whether or not radiation in low doses causes malignancies may be viewed from three perspectives:

a. First, it is a problem that should be clarified for its own sake, because of our increasing exposure to low levels of radiation in the form of medical diagnosis and therapy and numerous commercial

* Assigned to the University of Cincinnati Radioisotope Laboratory by the United States Public Health Service, Division of Radiological Health.

uses which lead to occupational exposures in minimal doses. In addition, a contaminated atmosphere is a reality which cannot be wished away by any amount of good will. A question of great social and economic importance in this cold war period is, "Just how much fallout radiation is still safe?"

b. Investigations into the effects of low dose radiation are of considerable methodological importance. Of all problems in so-called "chronic disease epidemiology," the parameters with which the effects of an agent over a long time period can be investigated can be defined here with some rigor. The dates of radiation procedure, reasons for radiation, and some information on physical status of the patient can be established with some certainty. Doses and sites of radiation can be approximated at least sufficiently well in most cases so as to make comparisons between patients possible. Populations of radiated individuals can be defined easily, these populations are large enough to permit testing of relevant hypotheses, and finally, the individuals in such populations are located with comparative ease. Thus, the low dose radiation problem lends itself ideally to the development of adequate epidemiological tools and concepts. Two major problems that beset studies in this area can be defined clearly, namely, adequate sampling procedures from relevant *eligible* populations and the evaluation of effects of an agent on morbidity and mortality when time-lags are long enough to permit *competing risks* from other disease causes to operate. Eligibility and competing risk have been discussed at some length in recent literature (Berkson and Elveback, 1960; Sterling *et al.*, 1962; Yerushalmy and Palmer, 1959). In essence, eligibility is the crux of the sampling problem. Populations may be divided into sets which differ in their disease patterns. Individuals with certain disease patterns may also be more likely to expose themselves to an environmental characteristic (such as radiation for diagnosis or therapy). Thus the joint probabilities of finding individuals with both characteristic and disease may be different for different population subsets. In consequence, unless groups from the same eligibility subsets are compared, statistical hypotheses of group differences cannot be tested (Latourette and Hodges, 1959).* Competing risks rest on the fact that the sum of the probabilities of death from any disease at any age are unity. Thus decreases or increases in the probability of death from some causes increase or decrease the probabilities of death from other causes. Another confounding mechanism in similar fashion may occur because increased risk of morbidity from some diseases is accompanied by increased or decreased risks in other disease morbidity (Sterling *et al.*, 1962). The

*The basic analysis on which concepts of eligibility are based were presented most lucidly by Yerushalmy and Palmer (1959.) When we wrote our own critique on radiation epidemiology we were, unfortunately, unaware of the existence of this paper so that no references were made to it. As much as it is possible to apologize for one's ignorance we do so now.

development of adequate epidemiological tools hinges largely on the success with which the epidemiologist will meet the problems of eligibility and competing risk in investigation on low dose radiation effects.

c. The final issue, and the one that makes this topic of concern to the geneticist, is the extent to which the confounding effects in different eligibility sets are of hereditary origin. The different disease patterns referred to before and demonstrated below may be defined completely by heredity, they may represent faults in the genetic armamentarium which are made apparent by the right kinds of environmental stresses, or they may be solely normal responses to environmental events.

Before turning to the analysis of pertinent evidence, it may be well to review why low dose radiation is a separate problem from that of radiation in high doses. Most known toxic agents do not act in a simple regularly increasing or decreasing relation between dosage and physiological effect. Most agents in the environment seem to follow a simple law as to their harmlessness or necessary body burden, which may be stated thus:

> Excessive amounts of the agent coming in effective contact with body tissue will have deleterious effects on health and, if the dosage is sufficiently increased, will lead to irreversible effects and death. Deficiencies of the agent in the environment or insufficient effective contact with body tissue may have deleterious effects on health and, if their dose is sufficiently decreased, may lead to irreversible effects or death. In many instances there exists then a range of dosage making up an effective body burden which is necessary as well as harmless.

The above is true of a large number of physical and chemical substances found in man's environment. Deficiencies in copper and iron are as harmful ultimately as are deficiencies in oxygen; and excessive amounts of oxygen are as deadly as excessive amounts of copper and iron. As a consequence one can speak of levels which are necessary to health and of neighboring levels which are harmless deviations from requirements.

Whether or not the law of necessary and/or harmless body burden applies to radiation has not been established. However, because the law appears to hold for an increasing number of substances that occur naturally in the environment, one cannot disregard its possible application. In general, if a substance is shown to be harmful in high doses it is therefore not necessarily a demonstrated hazard in low doses. As such the low dose radiation problem stands on its own evidence.

Investigations on the effects of low dose radiation have been of two basic types. The first type of study employed large samples of individuals who had been treated for benign conditions in the past or had been exposed to radiation during diagnosis. These individuals (or their offspring) were followed to observe the incidence of neoplastic disease. This incidence was then compared to the incidence of neoplastic disease among siblings or playmates of the study group or with that of the general population. The second approach assembled the radiation records of individuals dying with tumor or leukemia. The incidence of radiation in this group was compared to that of a matched control population. In either case, the crucial assumption was made implicitly, that the difference between radiated and control groups lay solely in the event of radiation.

Before coming to closer grip with our problem, the reason for the widespread use of irradiation in populations presently available for study must be realized and understood. Prior to the common use of sulfonomides, beginning about 1942, and the introduction of antibiotics beginning about 1945, treatment of infections in children was difficult and in many cases expectant. Radiation therapy was commonly and widely utilized and was considered to produce satisfactory clinical responses in many circumstances. Diagnoses were comparatively vague and gunshot use of radiation preceded the similar use of antibiotics in the recent decades. Radiation was used freely also for such acute infections as acute cervical adenitis, tonsillitis, mastoiditis, pneumonitis, bronchitis and for certain chronic diseases as tuberculous adenitis, hypertrophied tonsils and adenoids, acne vulgaris and other dermatoses.

In this group, one so called disease turned out to be especially important for subsequent studies. Many diseases of the head, neck and chest were attributed to "enlarged thymus," an alleged entity which, in retrospect, was completely misunderstood by physicians. The association of an apparently enlarged thymus gland in infants and children dying suddenly gave rise to the diagnosis of status thymolymphaticus and the development of radiography produced the concept of enlarged thymus from roentgenograms of the chest taken at short target film positions (about 30″) with the child in the supine position which maximized the size of the gland. All of these suppositions totally ignored the observations of Hammar (1906) and later Boyd (1932) which clearly showed that at least 50 percent of normal health children have thymus weights greater than the "accepted upper limit" of 20 gms. The rapid decrease of cell population and size of lymphoid structures with many types of stress was not recognized nor was the equally rapid recovery of thymic weight appreciated. Thus children dying after chronic illness were found to

have acceptable small thymus glands whereas those who died suddenly in good nutritional status demonstrated "enlarged" glands which satisfied medical curiosity as an apparent cause of death and prevented further search for other factors.

Thus the study population consists of individuals who were irradiated because of a variety of illnesses and who were followed in time to observe the incidence of yet another specific disease. Such a population does not constitute a random sample by any stretch of the imagination nor can one compare such a group to individuals who did not have a similar history of disease. Yet, most investigators proceeded as if the radiologist had in the past snatched his patients at random from the crowd that passed by his office. It is difficult to understand why so many studies were done which contain this very flaw in their design. For instance, table 1 lists the major populations on which the relations between leukemia and ionizing radiation have been studied. With exception of radiologists and victims of atomic warfare (exposed in both cases to extremely high doses) patients come from populations which suffered from disease in one form or another. As may be expected, the outcome of these investigations resulted in anything but enlightenment. We can see this very clearly when we review three statements which relate radiation in small doses to the subsequent development of cancer:

1. Irradiation of head, neck, and chest (for benign conditions) leads to the development of malignant neoplasms;
2. Irradiation of ankylosing spondylitis leads to the subsequent development of leukemia;
3. Prenatal irradiation leads to the subsequent development of leukemia.

MALIGNANT NEOPLASMS AND IRRADIATION TO THE HEAD, NECK, AND CHEST REGION

Many studies have investigated the relationships between radiation therapy given to the head, neck, and chest for begnign conditions and the subsequent development of neoplasms. Table 2 lists these studies and summarizes them. Only three studies give sufficient controls to merit discussion here. Two of these, Simpson *et al.* (1955, 1957) and Saenger *et al.* (1960), were similar in design and execution. Individuals were selected for radiation therapy (the characteristic) because of the presence of certain diseases or symptom complexes of the upper or lower respiratory tracts. These conditions at the time of radiation were diagnosed as enlarged thymus, acute cervical adenitis, hypertrophied tonsils and adenoids, bronchitis, and bronchopneumonia, tuberculous adenitis, etc. The subsequent rate of malignancies in this population was compared to that of their siblings, which formed the control group.

TABLE 1.—*Major populations in which the relationship of ionizing radiation and leukemia has been studied*

	Age Group	Reason for Exposure	Source and Type of Radiation	Dose Range	Time Factor	Area Irradiated	Antecendent Diseases
I	Prenatal	Medical diagnosis	Diagnostic x-ray	Less than 2	Usually single	Whole body	Obstetrical difficulty of varying etiology.
II	Children	Medical therapy	Therapeutic x-ray	100–3000	Single, fractionated	Head, neck and chest area.	Enlarged thymus, cervical adeni, other benign conditions.
II	Children and adults.	Atomic weapon detonation.	Nuclear fission gamma-neutron.	Up to 1200+ over a very short time interval.	Single	Whole or partial body.	Possibly poor nutrition "ravages of war."
IV	Adult	Occupational (Radiologists).	Diagnostic x-ray gamma (radium).	5–100+ per year	Daily discontinuous, protracted.	Whole or partial body.	None.
V	Adult	Medical therapy	Therapeutic x-ray	112–3535*	Single, fractionated	Spine	Ankylosing spondylitis.
VI	Adult	Medical therapy	I-131 beta-gamma	5–3000+	Continuous low dose rate protracted.	Blood and whole body.	Thyroid disease.
VII	Adult	Medical therapy	P-32 beta	100–450+	Continuous low dose rate protracted.	Bone marrow and whole body.	Polycythemia vera.

*Mean dose to spinal marrow.

TABLE 2.—*Major retrospective-prospective studies of the relationship between therapeutic irradiation of benign conditions of the head and neck and subsequent neoplasm*

Study	Cases in irradiation/cases in controls				Total pts. irradiation	Controls	Population irradiation	Reason for irradiation	Dose range "r"	Incidence of thyroid CA per 100,000 pts.	Incidence of leukemia and lymphoma per 100,000 pts.
	Thyroid	Adenoma thyroid	Leukemia hymphoma	Other cancer							
Simpson *et al.*, 1955, 1957, 1961.	11/0	9/1	9/0	4/6	2,081	2,722 sibs	Children	Thymus	100–600	528/0	432/0
Latourette & Hodges, 1959.	1/—	0/—	2/—	1/—	867	0	Children	Thymus	50–600	115/—	231/—
Snegiriff, 1959	0/0	2/0	0/0	0/0	148	185 nonirradiated	Children	Thymus	100–1,800	0/0	0/0
Moloney, 1959	2/—	5/—	0/—	0/—	125	0	Children	Thymus	50–17,150	1.600/—	0/—
Murray *et al.*, 1959	—/—	—/—	2/0.28	—/—	1,148	New York State Statistics.	Children	Benign condition of head and neck excluding thymus.	Not given		174/—
Conti *et al.*, 1960	0/0	0/0	0/3	0/0	1,564	2,923 sibs	Newborn	Routine thymus	75–450	0/0	
Saenger *et al.*, 1960	10/0	7/—	1/3	6/3	1,644	3,777 sibs	Children	Benign condition about head and neck.	50–1,200	609/0	0/102 61/79
Newman, 1960	0/1	9/—	0/—	0/—	32	0	Children	Thymus	180–1,540	0/—	
Stasek, 1962	1/—	0/—	0/—	0/—	52	0	Children	Cervical adenitis	100–420	1,923/—	0/—
Hanford *et al.*, 1962	8/0.1	—/—	—/—	13/12.50	295	Conn. Dept. Health Statistics.	Children and adults.	Thymus, hyperthyroid, TB adenitis.	500–2,000	1,920/—	0/— —/—

Assuming that patients with enlarged thymus were symptomatic, albeit lacking a more precise diagnosis this assumption seemed quite tenable, these patients could not be classed as having been randomly picked from a normal population. To test and select among the alternatives,

1. that irradiated children have the same rate of malignancies as their siblings,
2. that irradiated children have a greater frequency of malignancies,

one must assume that either hypothesis or alternative could be possible. However, there is the possibility that children referred for radiation therapy, although for benign conditions, were more eligible to develop malignancies eventually than were the controls.

Evidence against the eligibility hypothesis was found in the study by Saenger *et al.* (1960). Index cases not only had a greater incidence of neoplastic but of all other disabling diseases. The greatest discrepancy was in circulatory diseases and index cases had, on the average, 2.7 times as many disabling diseases as did the controls (table 3). Since the index cases had a much higher rate of diseases other than malignancies, it may be reasonable to assume that they could have had a higher spontaneous incidence of malignancies. At any rate, a statistical test between the rates of index cases and controls is not justified.

TABLE 3.—*Ratios of patient/sibling incidence of disabling diseases*

(adjusted for contributing life-years)
(Saenger *et al.*, Radiology, 1960, Table XI)

Category of disabling diseases	Siblings per 100,000	Patients per 100,000	Patients/ siblings
Circulatory diseases	21.9	115.5	4.27
Malignant neoplasms	16.1	59.4	3.7
Respiratory diseases	31.1	115.5	3.59
Nervous diseases	29.2	95.7	3.28
Blood and blood forming diseases	10.2	33.0	3.23
Mental diseases	30.6	82.5	2.69
Infections	17.5	42.9	2.44
Genitourinary diseases	24.8	46.2	1.86
Digestive diseases	19.0	23.1	1.22
Allergic, endocrine, metabolic, and nutritional diseases	93.4	99.0	1.06

Mean ratio 2.74±1.31

It seems not unlikely that the same factors were operating in the study by Simpson *et al.* (1955, 1957). This study is alike in all comparable dimensions to that of Saenger *et al.* (1960).

The third study is that of Conti *et al.* (1960) in which a group of 1,564 newborn received thymic irradiation (table 2). The first subgroup of 224 who were essentially asymptomatic but were diagnosed roentgenologically as having enlarged thymic shadows, were treated, and later a second subgroup of 1,340 infants born subsequently were so treated. This population is unique in several respects.

All patients of an apparently normal population were included in the study thus satisfying the criterion of eligibility. The controls were the unirradiated siblings as in the other studies. The only selection factor was that the patients were born in a particular hospital—they were for the greater part (87.2 percent) not selected because of the presence of any symptom, disease or laboratory finding. Thus it would seem likely that both patient and sibling had potentially an equal chance of developing subsequent neoplasm save for the presence of radiation (the characteristic) in the patient group. No carcinomas of the thyroid were found in either group; three leukemias were found in the siblings but none were found in the patient group. In comparing the findings of this study to the two others, one must remember that the irradiation portals were significantly smaller and perhaps that the dose to the thyroid itself was significantly less.

How is one to summarize these three studies? Can one say that because Simpson *et al.*, and Saenger *et al.*, found an increasing neoplasia that these conditions were caused by radiation? Can one say that because Saenger *et al.*, found an increase in all diseases that one dealt here with an insult falling, as it were, on already prepared soil? Can one say that because of findings by Saenger *et al.*, and because of the lack of findings in the study of Conti *et al.* (1960) which had a more truly random population that radiation had no effect at all? Clearly none of the studies offers a clear cut answer to these questions.

IRRADIATION OF ANKYLOSING SPONDYLITIS AND THE DEVELOPMENT OF LEUKEMIA

Court Brown and Doll (1957) studied the incidence of leukemia in a group of patients with ankylosing spondylitis who had been given radiation therapy for symptomatic relief. A summary of the results of that study with respect to observed death from leukemia is given in table 4. In addition, the authors found some evidence for a dose response relationship relating relevant cases of leukemia to man years at risk following various levels of mean spinal marrow dose for the group of patients receiving only spinal irradiation and under some other special conditions.

TABLE 4.—*Leukemia following irradiation for ankylosing spondylitis (Court Brown and Doll, 1957)*

Total patients	13, 352
Deaths due to leukemia	28
Expected number of deaths (calculated from Brit. Vital Stat.)	2. 9

As in two of the previously cited studies of irradiation in children, the eligibility hypothesis is in doubt. Is there a greater probability of individuals with ankylosing spondylitis to develop leukemia than for individuals not suffering from this disease? Although an association between ankylosing spondylitis and leukemia was discussed as a

possibility, the authors state that available data at that time do not suggest a simple association of the two diseases in the absence of radiation (p. 21). Similarly, various drugs did not seem implicated as associated leukemogens. The suitable control population, *i.e.*, a group of individuals with ankylosing spondylitis who had not received radiation, was not available for comparison.

Abbatt and Lea (1958) reported in 1958 a group of 661 patients with leukemia and not treated with irradiation or sulfa drugs, of whom 54 or 8.2 percent, had some type of rheumatic disease and 37 or 5.6 percent had confirmed rheumatic disease. Of a control series of 1,378 patients with diseases other than leukemia, 51 or 3.7 percent had some type of rheumatic disease and 20 or 1.5 percent had confirmed rheumatic disease (table 5). There appears to be here a relation between untreated rheumatism and leukemia; consequently, a test of increase in leukemia rate as a function of irradiation is not feasible.

TABLE 5.—*Summary of data of Abbatt & Lea (1958)*

	All cases of rheumatic disease (1)	Confirmed cases of rheumatic disease (2)	No rheumatic disease (3)	Subtotal (1+3) (4)	Excluded because of radiation (5)	Grand total (4+5) (6)
Patients with leukemia	54	37	607	661	18	679
Controls	51	20	1,327	----------	----------	1,378
Percent of leukemia patients with associated spondylitis.	$\frac{54}{661} \times 100 = 8.2\%$					
	$\frac{37}{661} \times 100 = 5.6\%$					
Percent of control patients with associated spondylitis.	$\frac{51}{1,378} \times 100 = 3.7\%$					
	$\frac{20}{1,378} \times 100 = 1.5\%$					

As before, one appears to have a choice of three different conclusions. The increased rate of leukemia for irradiated ankylosing spondylitis patients might be taken to indicate an association between leukemia and radiation therapy; the fact that rheumatic conditions and leukemia are associated would indicate that the increased incidence of leukemia in the irradiated group is spurious; finally, the somewhat tenuous dose response relationship might point to a case where an insult falls on fertile soil. A clear cut distinction is again not possible.

PRENATAL IRRADIATION AND THE SUBSEQUENT DEVELOPMENT OF LEUKEMIA IN CHILDREN

A third group of papers which have been given considerable attention are given in table 6. The largest number of cases and the best

TABLE 6.—*Retrospective studies of the relationship between prenatal irradiation and leukemia*

Study	Number cases irrad./Total number cases (percent)			Selection of controls	Source of information
	Leukemia	Other cancer	Controls		
Stewart *et al.*, 1956, 1958, 1961.	96/780 (12.3)	115/858 (13.4)	117/1638 (7.1)	Matched for age, sex and location	Personal interview.
Kjeldsberg, 1957	5/55 (9.1)		8/55 (14.5)	"Healthy children"	
Kaplan, 1958	a) 37/150 (24.7) b) 34/125 (27.2).		a) 24/150 (16.0) b) 27/125 (21.6).	a) Closest sibling b) Most habitual playmate.	Personal interview.
Polhemus & Koch, 1959	72/251 (28.7)		58/251 (23.1)	Children from same hospital with other selected conditions.	Questionnaires by mail.
Ford *et al.*, 1959	White 20/70 (28.6) Colored 1/8 (12.5).		White 48/247 (19.4) Colored 8/59 (13.6).	Children dying under 10 years of causes other than cancer.	Hospital and physician records.
Murray *et al.*, 1959	3/65 (4.6)		a) 3/65 (4.6) b) 7/93 (7.5). c) 2/82 (2.4).	a) Children dying under 20 years of causes other than cancer. b) Siblings of leukemia children. c) Siblings of (a).	Personal interview.
Lewis, 1960	1/8 (12.5)		(25.3)	Incidence of prenatal irradiation in all children born in the same hospital.	Hospital records.
MacMahon, 1962	47/304 (15.5)	38/252 (15.1)	770/7242 (10.6)	Incidence of prenatal irradiation in a 1-percent sample of all births in same hospitals.	Hospital records.
Ranges	4.6=28.7	3.4=15.1	7.1=25.3.		

controlled series is that of Stewart *et al.* The original communication (1956) reported that in a group of 547 children dying of malignant disease, the mothers of 85 patients had received diagnostic X-ray studies (usually for pelvimetry) when the patient was *in utero* as compared to the mothers of 45 patients in the control group of 547. Thus an estimated dose of 2.5r was thought to be cancerogenic. In a subsequent paper, Stewart *et al.* (1958) reported on a total of 1,416 cancers in children of which 677 were due to leukemia and 739 to other cancers. In this paper the ratio of maternal abdominal diagnostic studies during the relevant pregnancies was 178 as compared to 93 in the controls (table 7).

TABLE 7.—*Number of mothers reporting X-ray examination of different sites (abdomen and others) in three periods*

(Stewart *et al.*, Brit. M. J., 1958, Table II)

Period	X-ray examinations		
	Abdominal	Other	Any
Before marriage	44/26=1.69	335/275=1.22	361/296=1.22
Between marriage and relevant conception	109/121=0.90	213/184=1.16	304/285=1.07
During relevant pregnancy	178/93=1.91	117/100=1.17	273/184=1.48
Any period	296/215=1.38	531/456=1.16	692/593=1.17

The technique of this study was to obtain the names and medical histories of all children who died with leukemia or other tumors before their 10th birthday in the years 1953–55 in England and Wales. The controls were then matched for age, sex, and locality but were otherwise selected at random. Again the pattern of comparison of a selected case population vs. a random control population suggests that the eligibility hypothesis is in question.

Eligibility and heredity may well be related. In the present instance, one may ask if mothers who require more medical attention are at the same time of greater disease potential and might not pass on this increased tendency to their offspring? If the latter possibility is taken seriously, as indeed it should be, some evidence must be obtained that this hereditary effect does not make itself felt before a meaningful test between malignant and controls is possible.

Some scant information is given by Stewart *et al.* (1958) on this point. Table 8 summarizes the incidence of childhood diseases for malignant and control children. If one subtracts the common infectious childhood illnesses to which most children are prone, one obtains a ratio of diseases between those with and those without malignancies of 1 to 5. The justification of computing such a ratio is not too clear since children with malignancies may also show less resistance to disease in general. Somewhat more helpful is the comparison of congenital diseases. Here the ratio between malignant and control cases equals 1 to 6. Again there is an increased incidence of congenital mal-

TABLE 8.—*Ratios of patient/control incidence of childhood disease*
(Stewart *et al.*, Brit. M. J. 1958, Table XX)

	Totals of dated illnesses		
	Survey children (all malignant diseases)	Controls	Ratio patient/control
Bronchitis and broncho-pneumonia	133	92	
Acute tonsillitis and otitis media	79	68	
Other infections	153	101	
Other illnesses	58	24	
Acute exanthemata	1,181	1,241	
Grand total	1,604	1,526	
Grand total less exanthemata	433	285	1.5
Congenital defects	75	46	1.6

formations in the malignant group that compares to the ratio of prenatal radiation. However, again this comparison is not very convincing. First, such malformations could possibly have been caused by radiation and secondly, many of the congenital cases are of doubtful importance; the real difference is that between mongols, there being 18 in the malignancies and none in the controls.

Court Brown and Doll (1960) reported a similar study conducted somewhat differently. They studied the incidence of leukemia mortality of children exposed to antenatal radiography in 4 London and 4 Edinburgh hospitals from 1 Jan. 1945 to 31 Dec. 1956. The relevant case records of 39,166 live-born children in England, Wales and Scotland were then reviewed. In those children not receiving radiation, 9 leukemia deaths were observed, 10.5 being expected. For those in whom pelvimetry was performed, 4 leukemia deaths were found, 5.4 being expected. For those in whom a single abdominal exposure was carried out, 5 leukemias were found, 4.3 being expected.

In this study the eligibility hypothesis is under question for the same reason as it was challenged by the Stewart study. Again one would be led to assume that mothers given x-radiation for diagnostic reasons on the whole represent a sicker population than mothers not requiring such radiation. Unfortunately the record of other diseases is not given by Court Brown and Doll.

The contribution of competing risks to the low dose radiation problem has not been studied intensively so far. One interesting aspect of this contribution emerged from the study of Saenger *et al.* (1960). In this study, irradiated patients showed a much higher incidence of all diseases than did their nonirradiated sibs. It was postulated, therefore, that two separate populations were compared, one with a higher and the other with a lower incidence of disease in general. However, if this were the case, one could also expect that the group with a high incidence of morbidity would show up among the siblings in a peculiar fashion. If indeed there were these separate population sets, one could expect that some individuals in the high morbidity set were so ill that they died early without prolonged illness and thus did

not come to the radiologist at all. Since such individuals would not have radiation exposure in their history they would be classified among the siblings. Thus one would expect to observe a large number of excess deaths among the siblings during their early years of life. This turned out to be the case indeed. Up to age 14, siblings died at a much higher rate than did either irradiated patients or the population of the U.S. or the state of Ohio.

CONCLUSION

In problems that go beyond the lifespan of the commonly used laboratory animal and may even last for more than the useful lifespan of the average investigator, one has to turn to techniques outside the laboratory. As unpalatable as it may be to some to rely on non-experimental proofs for biological problems, no other solution appears to be in sight. This approach is doubtless true for studies that wish to investigate the relation between small dose radiation and malignancies. Such studies are necessary not only to establish a clear and present danger where it exists, but also to retain radiation as a therapeutic and diagnostic tool in the face of mounting hysteria based on questionable evidence.

Awareness of the sources of error, subsumed here under the concept of eligibility, should make it possible to design studies that will give a satisfactory answer to the problem of hazards associated with small dose radiation in most cases. It may be necessary to select samples over a wider range of clearly defined population categories and to apply more sophisticated mathematical analyses to evaluate the observed results. For example, a satisfactory design for an investigation asking questions as that of Simpson *et al.* (1955, 1957) or Saenger *et al.* (1960) should sample from two sets of index cases and their siblings.

INDEX I: Patients irradiated for benign conditions.

CONTROL I: Siblings and parents of cases in INDEX I.

INDEX II: Patients treated for benign conditions similar to those in INDEX I, but not irradiated.

CONTROL II: Siblings and parents of cases in INDEX II.

Information collected would be similar to that in Saenger *et al.* (*i.e.*, including malignancies as well as a wider range of other, clearly defined diseases). The analysis will adjust the observed rate of diseases in INDEX II for:

a. the differences in malignancies and other diseases between CONTROL I and II siblings,

b. the differences in probability of developing malignancies and other diseases given differences in these diseases in the parental stocks of CONTROLS I and II,

c. differences in relative probabilities (or prevalences) of different disease categories between CONTROL I and II.

The crucial comparison would then be the observed incidence of malignancies (both general and in the areas irradiated) in INDEX I with the expectation based on the adjusted incidence of INDEX II. Such a study will answer the question if radiation may be considered more or less traumatic for a predisposed population than would be the case for other methods of treatment.

Such a design still leaves unanswered the very interesting question of the possible contribution of hereditary factors. Before engaging in an obvious familial study, one could, by adding one more group, check on the hypothesis that the population from which individuals are selected for radiation treatment, has a different general tendency toward morbidity and mortality. To accomplish this one would add another category of index and control cases:

INDEX III: Individuals who were not treated for benign conditions matched socioeconomically to INDEX I and II cases.

CONTROL III: Siblings and parents of INDEX III cases.

By comparing CONTROL III to the other controls one could establish the presence (albeit not the nature) of hereditary factor. Such a study as described here is now being executed by the authors. Results will be available in 3 years and should clarify the problems stated here.

The one definite conclusion that emerges from our review is that problems of chronic disease epidemiology such as the effects of low dose radiation are highly statistical as well as biological. Statistics has been called rightly the effort of making common sense rigorous. In this field in which so many statisticians and biologists have labored together, neither have shown much of rigor or common sense. What may be necessary is that both attempt to learn enough about each other's discipline so that they may investigate adequately what appears to be one of the pressing problems in today's medical science.

SUMMARY

The field of radiology appears to serve as a focus for reanalysis of epidemiological tools and concepts because, despite diligent applications of approved procedures, a number of epidemiological investigations have failed to answer questions on the possible hazards arising from relatively small radiation doses used in diagnoses and during some treatment. Two major problems that beset studies in this area can be defined clearly, namely, eligibility and competing risks. When studies purporting to link small dose radiation to subsequent development of neoplasia are analyzed critically in the

light of eligibility criteria, it becomes quite apparent that they do not at all establish such an association. Rather a suspicion could be entertained that radiation and the subsequent development of tumors or leukemias are due to a common factor which may be rooted in inherited disease tendencies. A study is described that will, hopefully, solve the problem of the effect of small dose radiation and also test the hypothesis of inherited disease factors.

ACKNOWLEDGMENTS

The authors are beholden to the few critics in the modern scientific area; especially do we acknowledge the influence of Berkson, Fisher, Hogben, and Yerushalmy on our work.

REFERENCES

1. Abbatt, J. D., and Lea, A. J., 1958: Leukaemogens. *Lancet* 2: 880–883.
2. Berkson, J., 1959: Statistical Investigation of Smoking and Cancer of Lung. *Proc. Staff Meet. Mayo Clin.* 34: 206–224a.
3. Berkson, J., and Elveback, L., 1960: Competing Exponential Risks, with Particular Reference to the Study of Smoking and Lung Cancer. *J.A.S.A.* 55: 415.
4. Boyd, E., 1932: The Weight of the Thymus Gland in Health and Disease. *Am. J. Dis. Child.* 43: 1162–1214.
5. Conti, E. A., Patton, G. B., Conti, J. E., and Hempelmann, L. H., 1960: Present Health of Children Given X-ray Treatment to the Anterior Mediastinum in Infancy. *Radiology* 74: 386–391.
6. Court Brown, W. M., and Doll, R., 1957: Leukaemia and Aplastic Anaemia in Patients Irradiated for Ankylosing Spondylitis. *Medical Research Council Special Report Series* No. 259. London, England. Her Majesty's Stationery Office.
7. Court Brown, W. M., and Doll, R., 1960: Prospective Study of Leukaemia Mortality of Children Exposed to Antenatal Diagnostic Radiography; Preliminary Report. *Proc. Roy. Soc. Med.* 53: 761–762.
8. Ford, D. D., Paterson, J. C. S., and Trueting, W. L., 1959: Fatal Exposure to Diagnostic X-rays and Leukemia and Other Malignant Disease in Childhood. *J. Nat. Cancer Inst.* 2: 1093–1104.
9. Hammar, J. A., 1906: On the Weight, Involution and Persistence of the Thymus in Post-Fetal Life in the Human. *Arch. f. Anat. u. Entwcklngsgesch Suppl.* p. 91.
10. Hanford, J. M., Quimby, E. H., and Frantz, V. K., 1962: Cancer Arising Many Years after Radiation Therapy. Incidence after Irradiation of Benign Lesions in the Neck. *JAMA* 181: 404–410.
11. Kap'an, H. S., 1958: The Evaluation of the Somatic and Genetic Hazards of the Medical Use of Radiation. *Am. J. Roent.* 80: 696–706.
12. Kjeldsberg, H. T., 1957: Radioaktiv Bestrailing og Leukemefrekvens hos barn. *Tidsskr. Norske Laegeforen,* 77: 1052–1053. Quoted by reference (Wold and Thoma).
13. Latourette, H. B., and Hodges, F. J., 1959: Evidence of Neoplasia after Irradiation of Thymic Region. *Am. J. Roent.* 82: 667–677.

14. Lewis, T. L. T., 1960: Leukemia in Childhood after Antenatal Exposure to X-rays. *Brit. M. J.* 2: 1551–1552.
15. MacMahon, B., 1962: Prenatal X-ray Exposure and Childhood Cancer. *J. Nat. Cancer Inst.* 28: 1173–1191.
16. Moloney, W. C., 1959: (In) *Radiation Biology and Cancer.* Austin, Texas, University of Texas Press, p. 344.
17. Murray, R., Heckel, P., and Hempelmann, L. H., 1959: Leukemia in Children Exposed to Ionizing Radiation. *New England J. Med.* 261: 585–589.
18. Murray, R., and Hempelmann, L. H., 1961: A Review of Tumor Evidence in Children Irradiated for Benign Conditions. (In) *Radioactivity in Man,* G. R. Mencely, ed., Springfield, Illinois, Charles C. Thomas, p. 282.
19. Newman, G. G. H., 1960: Long-term followup of Thirty-two Patients Irradiated for Thymic Enlargement in Infancy. *Brit. M. J.* 1: 34–36.
20. Polhemus, D. W., and Koch, R., 1959: Leukemia and Medical Radiation. *Pediatrics* 23: 453–461.
21. Saenger, E. L., Silverman, F. N., Sterling, T. D., and Turner, M. E., 1960: Neoplasia following Therapeutic Irradiation for Benign Conditions in Childhood. *Radiology* 74: 889–904.
22. Simpson, C. L., Hempelmann, L. H., and Fuller, L. M., 1955: Neoplasia in Children Treated with X-rays in Infancy for Thymic Enlargement. *Radiology* 64: 840–845.
23. Simpson, C. L., and Hempelmann, L. H., 1957: The Association of Tumors and Roentgen Ray Treatment of the Thorax in Infancy. *Cancer* 10: 42–56.
24. Snegiriff, L. S., 1959: The Elusiveness of Neoplasia following Roentgen Therapy for Thymus Enlargement in Childhood. *Radiology* 72: 508–517.
25. Stasek, V. Jakoubkova, J., Brachfield, K., Kolar, J., and Tichy, S.. 1962: Latent Reaction in Children Treated with Small Doses of X-rays. *Strahlentherapie* 117: 293–299.
26. Sterling, T. D., Saenger, E. L., and Phair, J. J., 1962: Epidemiology. *Cancer* 15: 489–503.
27. Stewart, A., Webb. J., Giles, D., and Hewitt, D., 1956: Malignant Disease in Childhood and Diagnostic Irradiation in Utero. *Lancet* 2: 447.
28. Stewart, A., Webb, J., and Hewitt, D., 1958: A Survey of Childhood Malignancies. *Brit. Med. J.* 1: 1495–1508.
29. Stewart, A., 1961: Aetiology of Childhood Malignancies Congenitally Determined Leukemias. *Brit. Med. J.* 1: 452–460.
30. Yerushalmy, J., and Palmer, C. E., 1959: On the Methodology of Investigations of Etiologic factors in Chronic Diseases. *J. Chronic Dis.* 10: 27–40.

Congenital Malformations and Cancer in the Sibships of Leukemic Children: Further Considerations*

By Robert W. Miller, M.D., Dr. P. H.,
National Cancer Institute, National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare, Bethesda, Md.

In the study of the effects of chromosomal aberrations on health there has been a most effective interaction between epidemiologic and laboratory research. Epidemiologic studies showed that leukemia occurs much more frequently than expected in mongolism (Krivit and Good, 1957) and after sufficient exposure to ionizing radiation (Folley, Borges, and Yamawaki, 1952; and Brill, Tomonaga and Heyssel, 1962). Laboratory research later revealed that common to all three events are visible chromosomal defects (Lejeune, Gautier, and Turpin, 1959; Ford, Jacobs and Lajtha, 1958; Baikie *et al.*, 1959; Nowell and Hungerford, 1960; and Bender and Gooch, 1962). Whether the defect is inherited, as in mongolism, or acquired, as from radiation exposure, it is correlated with a higher risk of leukemia.

Epidemiologic research has shown that the older the woman, the more likely her pregnancy will terminate in miscarriage (Germany, Statistiches Reichsamt, 1910; and Shapiro, 1962), or in the birth of a mongoloid infant (Shuttleworth, 1909; and Penrose, 1934) or a normal infant who will develop leukemia later in childhood (Stewart, Webb, and Hewitt, 1958). Laboratory studies have recently indicated why mongolism and miscarriages increase with maternal age: most mongolism and some miscarriages are the result of meiotic nondisjunction, a chromosomal aberration which apparently increases in frequency with the age of the ova (reviewed by Ford, 1961). Perhaps a similar mechanism will explain the association between childhood leukemia and maternal age.

*Based on a previous publication by the author in collaboration with the project directors of the National Cooperative Leukemia Survey (Miller, 1963).

Epidemiologic observations have furthermore shown that mongols aggregate in sibships more often than can be attributed to chance (Penrose, 1938). Cytogenetic studies have identified several chromosomal mechanisms which are responsible (reviewed by Hamerton, 1962).

There is evidence, then, that specific chromosomal aberrations may be associated with several disorders and may be transmitted from a parent to more than one child. Epidemiologic studies are needed to determine if these several disorders occur more often than usual among the sibs of mongols or of leukemic children; *i.e.*, sibs of propositi with known or suspected chromosomal aberrations.

METHODS

Available for analysis were the data from questionnaires obtained between 1958 and 1961 through the cooperation of 11 medical centers and one State Health Department which participated in the National Cooperative Leukemia Survey (NCLS).* Nurse-interviewers had obtained detailed personal and family histories on standard forms for 459 non-Negro children with leukemia and for neighborhood control children individually matched with the leukemic children by race, age, birth order, and family size. Also available were records for an additional 60 leukemic children without controls. Hospital and other medical records were reviewed for verification of the history wherever possible.

The data were not collected specifically for the use to which they are being put here. (The cytogenetic information on which this study was based was not yet known at the inception of the survey.) The history of cancer and congenital malformations among the sibs had to be gleaned from questions about each child in the family regarding hospitalization, X-ray examinations and other "physical defects." The events observed were limited to those occurring before the onset of leukemia in the propositus.

The usual respondent was the mother. Since she may have failed to report all events of interest to us, rates obtained by interview were compared with results from other studies, and, when possible, with U.S. national vital statistics. Medical records served to verify events disclosed by the respondent, but for information not disclosed there is no estimate available as to differences, if any, between cases and controls. One estimate of this difference is being made by the group at Roswell Park Memorial Institute which is checking the respondents' statements against hospital records for sibs in the Tri-State Survey.

*A parallel study, The Tri-State Survey, was carried out by the Roswell Park Memorial Institute, Johns Hopkins University and the University of Minnesota. The principal analysis of the data from both surveys is being made at the Roswell Park Memorial Institute.

Review of the NCLS questionnaires revealed that a few of the nurse-interviewers did not query the informants as closely as did their counterparts in other cities. However, the failure to elicit positive responses seemed to apply equally to cases and controls.

The diagnosis of leukemia was made by a competent hematologist. Criteria for classifying cell-type have not been established, and there were obvious differences between physicians in this regard. For example, stem-cell leukemia was generally a rare diagnosis except at one clinic where there was a compensatory deficiency in all other cell types. Despite these limitations, hospital and physicians' records provided diagnoses of better quality than those generally found on death certificates, and the histories recorded by the nurses were much more complete and uniform than those recorded in essay style on medical charts.

One should consider whether leukemic children ascertained primarily through university hospitals will be 1) representative of leukemic children throughout the nation, and 2) comparable with neighborhood controls. In particular, since study was made of major ailments in sibships, one must ask if the results were apt to be biased by selective referral of leukemic children whose families have been beset by such illnesses. We suspect that because leukemia is so rare in the general practice of pediatrics, and so tragic, the need for accurate diagnosis probably overrides other medical considerations in the practitioner's selection of a hospital for referral of his patient. As a more objective estimate of the validity of the data, the results were evaluated for consistency with the body of knowledge already available about leukemia and with the findings of similar epidemiologic research employing different sampling methods.

RESULTS IN PERSPECTIVE

The coexistence of mongolism and leukemia was once again evident in this study. Six cases were found among the 459 leukemic children with matched controls and in one among the 60 leukemic children without controls. Since there is a heavy mortality among mongols early in life, the frequency of mongolism in the general population drops from an estimated 1.4/1000 at birth (reviewed by Penrose, 1961) to 1.0/1000 at 10 years of age (Carter, 1958). Thus the proportion observed, 7 in 519 children, was significantly in excess of the number expected in the general population or the number observed (zero) among the 459 controls.

The leukemic children in the matched sample had 1,000 sibs, 4 of whom were mongols. A fifth mongol was conceived 8 months after the cutoff date (onset of leukemia in index child). All of the cases were documented by hospital records. There were no mongols among

the 956 sibs of controls. Had there been 5 cases among the 1,000 sibs of leukemic patients the difference from the 1.4 cases expected in the general population of livebirths would have been statistically significant at the 5 percent level. There is, of course, the possibility that not all mongols were reported by the informants. The occurrence of mongolism among sibs of leukemic children has not previously been studied epidemiologically, but the excess suggested by this survey is consistent with several recent reports of individual sibships in which one child was leukemic and one or more others were mongols (Stewart, 1961 [Table VI, Case 2]; Miller *et al.*, 1961; Buckton *et al.*, 1961).

Major congenital defects other than mongolism affected 4.9 percent of the leukemic children as compared with 2.0 percent of the matched controls, a difference just significant statistically at the 5 percent level, and not attributable to any specific malformation. The information concerning the leukemic children was available from recent medical records. For the controls, however, the information had first to be elicited in the interview history, after which it could be checked from hospital records. Thus, underreporting was more likely among controls than cases. It is difficult to compare major congenital defects found in one survey with those in another for, among other variables, there are dissimilarities between observers, in the criteria for classifying defects as major or minor, and in survey techniques. However, a rough comparison was made with the frequency of the same congenital defects found by a comprehensive clinical study of 3,570 Japanese children 5 to 11 years of age (Schull and Neel, unpublished data). Among the Japanese children, 2.2 percent had comparable defects, a value similar to the 2.0 percent for the NCLS control group.

In a survey of 677 leukemic children Stewart, Webb, and Hewitt (1958) found that the only congenital defects other than mongolism in excess of normal expectation were naevi (considered minor defects in the NCLS). The absence of a published listing of the other defects found by Stewart and her associates prevents further comparison with the results of the NCLS.

Children with major congenital defects are more likely than healthy children to be examined or treated radiologically. Since the leukemogenic effect of ionizing radiation is well established, one must distinguish between a direct association of congenital defects and leukemia and an indirect one in which the primary association is between radiation exposure and leukemia. Review of the X-ray procedures described in the questionnaires for the 22 children with major congenital defects revealed the following: there were 6 with no exposures; 9 with 1 to 5 film studies of the head, chest, or extremities; 2 with abdominal fluoroscopy for intestinal disease; and 1 each had received a) a series of 5 hip X-rays, b) whole-body fluoroscopy in a practitioner's office,

c) intravenous and retrograde pyelography, d) a series of 3 spine, 4 skull, and 8 chest films, and e) cardiac catheterization plus a barium enema. Except perhaps for the last-mentioned child, the doses received appear to have been low. In studies of persons exposed to radiation for medical reasons, it is generally difficult to separate the effects of radiation from those of the underlying disease. When the underlying disease is a congenital defect there may be good reason for a predisposition to leukemia, if, as in mongolism, there is chromosomal imbalance present.

Among the 1,000 sibs of leukemic children, 26 had major congenital defects other than mongolism, as compared with 24 among 956 sibs of controls, an insignificant difference. The lack of a difference is not surprising. Malformations other than mongolism associated with chromosomal aberrations are estimated by Ford (1961) to be about as frequent as mongolism, but are varied in clinical appearance. Thus, such malformations would not be recognized among the sibs of leukemic children by either their clinical appearance or their frequency (about 5, if equal to the number of mongols).

Miscarriages were more often reported by mothers of leukemic children than by mothers of control youngsters (204 of 1,210 pregnancies vs. 149 of 1,113 pregnancies; index children excluded). These results are similar to those of Stewart, Webb, and Hewitt (1958). However, since they found *no* excess history of miscarriage among mothers of children with cancers other than leukemia, it seems likely that there was a true difference in miscarriage experience among mothers of leukemic children rather than a difference in maternal recall. It is important to note that if the difference in pregnancy experience is real, some childhood leukemia may be attributable to inherited chromosomal defects which in another pregnancy may produce fetal death.

Excessive stillbirth, however, was not observed in the NCLS or in the survey by Stewart, Webb, and Hewitt. In the NCLS there were 35 childhood deaths among the sibs of leukemic children as compared with 20 among sibs of the matched controls. Contributing to the excess, which was not statistically significant at the 5 percent level, were diseases of all categories except accidents.

In the matched series of sibships of leukemic children, there were 5 pairs of sibs and one trio with cancer (4 leukemias, 1 undifferentiated malignancy of the urinary bladder, 1 generalized carcinomatosis (mesothelioma), and 1 unspecified congenital abdominal tumor). In comparison, among the sibs of controls there occurred one cancer, a brain tumor. In the unmatched series, identical twins at 4 years of age both developed leukemia with onset 5 months apart. The difference in cancer experience of the groups compared may be further increased when the families have been completed and all living children have reached the upper age level of childhood leukemia.

735-712 O—65——25

On the other hand, contributing to the difference already observed is one sibship in which the pair of leukemic childern both developed their disease within the period of ascertainment for the study, thus providing this sibship with two chances, instead of one, to enter the study (the same was true for the leukemic twins in the unmatched series).

The validity of the greater cancer occurrence among sibs of leukemic children is strengthened by the similar finding made in the survey of Stewart, Webb, and Hewitt (1958) and Stewart (1961). The cancer excess among sibs is in accord with the observations in the series of Steinberg (1960), who found one leukemia and one other cancer among 458 sibs of leukemic children, and in the series of Campbell *et al.* (1961) and Stewart (personal communication) who found one leukemic, a mongol, among 232 siblings of leukemic children. The last-mentioned series also included a step-sibling who had died of a pontine glioma. Apparently, as Campbell *et al.* pointed out, 1,000 or more sibs of leukemic children must be studied to detect their increased risk of cancer.

The occurrence of an array of diseases more common among cases than controls may be real or may be due to differences between the groups compared with respect to a) the method of selection for the study, b) the information disclosed by the respondents, or c) the attitude of the nurse-interviewers. Selection bias appears to be an improbable explanation for the differences observed in the NCLS because Stewart, Webb, and Hewitt obtained results similar in many respects using data from 86 percent of all childhood leukemia deaths in England and Wales in 1953–1955. Also, in the British study, the associations found pertained to leukemic children but not to children with cancers of other sites, suggesting that the differences detected were real and not attributable to memory bias. Furthermore, in the NCLS there were many disorders studied which were not more common among cases than controls. Included among these were minor congenital defects in the index child, stillbirths, major congenital defects other than mongolism among the sibs, and maternal disease during pregnancy. These observations also suggest that memory bias was not primarily responsible for the differences found with respect to other variables.

The excess of mongolism and cancers among the sibs of leukemic children as compared with their controls cannot be ascribed to differences in birth order or family size since the index children of both groups were matched on these variables. Despite this matching, the mothers of leukemic children were on the average one year older than the mothers of the controls (27.2 vs. 26.2 years). Maternal age, however, did not play a prominent role in the occurrence of mongolism or cancers among the sibs of leukemic children. In families having a mongoloid child, none of the mothers at the birth of the index *leukemic*

child was over 34 years of age (23, 26, 32, 34 and 34 years; average 29.8 years). When the *mongoloid* children were born the mothers were 21, 25, 30, 33, and 40 years old (average 29.8 years). In families having a child with cancer in addition to the index leukemic child, at the birth of the propositus maternal ages ranged from 25–38 years (average 32.3 years); at the birth of the sib with cancer, maternal ages ranged from 21–44 years (average 30.3 years).

DISCUSSION

The results when reviewed individually are consistent with observations from previously published studies, and when viewed collectively form a pattern which suggests that disorders associated with childhood leukemia have in common a chromosomal defect. The relationship between childhood leukemia and these disorders parallel those of mongolism, in which a cytogenetic abnormality is regularly found. Mongolism and childhood leukemia, in addition to occurring together more often than can be attributed to chance, have each been associated with a) other major congenital defects, b) increasing maternal age at the birth of the affected child, c) an excessive history of miscarriage, and d) familial aggregation.

The study reported here suggests that there is an excess of mongols among the sibs of leukemic children. It remains for other epidemiologic studies to determine if the converse is true; i.e., if among the sibs of mongols there is an excess of leukemia. It also remains to be determined what other specific diseases, including those of adult life, are more common among mongols than in the general population. The identification of such diseases will suggest, as for childhood leukemia, that the genesis of the disease is somehow linked to cytogenetic abnormality.

By epidemiologic study one can recognize the excessive concurrence of disorders in a single individual or among his sibs. This new information heightens the value of the family history, for when the clinician detects a cluster of characteristics which may be related to a chromosomal aberration, cytogenetic study is indicated both as an aid to genetic counselling and for research to define the morphology and influence of the abnormal chromosome pattern. And, to complete the cycle, clinical and laboratory findings are likely to suggest new epidemiologic research, as they did for the study described here.

ACKNOWLEDGMENTS

I am indebted to the following project directors and the National Cancer Institute staff who planned the study and supervised the collection of data: Stella Booth, M.D., Trenton; Benjamin E. Carroll,

M.A., Bethesda; John H. Githens, Jr., M.D., Denver; Dankward Kodlin, M.D., M.P.H., Pittsburgh; M. Eugene Lahey, M.D., Salt Lake City; Mark H. Lepper, M.D., Chicago; W. Harding leRiche, M.D., M.P.H., Toronto; Arthur E. McElfresh, M.D., Philadelphia; Miriam D. Manning, M.D., M.P.H., Boston; Kirk T. Mosley, M.D., Dr. P. H., Oklahoma City; Edwin E. Osgood, M.D., Portland; Donald C. Smith, M.D., D.P.H., Ann Arbor; and Franklin H. Top, M.D., M.P.H., Iowa City.

The data collection was supported in part by grants from the National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare.

REFERENCES

1. Baikie, A. G., Court Brown, W. M., Jacobs, P. A., and Milne, J. S. 1959. Chromosome studies in human leukemia. *Lancet* 2: 425–428.
2. Bender, M. A., and Gooch, P. C. 1962. Persistent chromosome aberrations in irradiated human subjects. *Radiation Research* 16: 44–53.
3. Brill, A. B., Tomonaga, M., and Heyssel, R. M. 1962. Leukemia in man following exposure to ionizing radiation: a summary of findings in Hiroshima and Nagasaki, and comparison with other human experience. *Ann. Int. Med.* 56: 590–609.
4. Buckton, K. E., Harnden, D. G., Baikie, A. G., and Woods, G. E. 1961. Mongolism and leukemia in the same sibship. *Lancet* 1: 171–172.
5. Campbell, A. C. P., Gaisford, W., Paterson, E., and Steward, J. K. 1961. Tumors in children: a survey carried out in the Manchester region. *Brit. Med. J.* 1: 448–452.
6. Carter, C. O. 1958. A life-table for mongols with the causes of death. *J. Ment. Defic. Res.* 2: 64–74.
7. Folley, J. H., Borges, W., and Yamawaki, T. 1952. Incidence of leukemia in survivors of atomic bomb in Hiroshima and Nagasaki, Japan. *Amer. J. Med.* 13: 311–321.
8. Ford, C. E., Jacobs, P. A., and Lajtha, L. G. 1958. Human somatic chromosomes. *Nature* 181: 1565–1568.
9. Ford, C. E. 1961. Human cytogenetics. *Brit. M. Bull.* 17: 179–183.
10. Germany, Statistiches Reichsamt. 1910. Krankheits- und Sterblishkeitsverhältnisse in der Ortskrankenkasse für Leipzig und Umgegend. Berlin. Cited in U.N. Dept. of Social Affairs. 1954. Fetal, Infant, and Early Childhood Mortality. Population Studies No. 13. United Nations, New York.
11. Hamerton, J. L. 1962. Cytogenetics of mongolism. In: *Chromosomes in Medicine*, J. L. Hamerton, ed., p. 140–183. London: The Medical Advisory Committee of the National Spastics Society in Association with Wm. Heinemann Ltd.
12. Krivit, W. and Good, R. A. 1957. Simultaneous occurrence of mongolism and leukemia. Report of a nationwide survey. *A.M.A.J. Dis. Child.* 94: 289–293.
13. Lejeune, J., Gautier, M., and Turpin, R. 1959. Les chromosomes humains en culture de tissus. *C.R. Acad. Sci. (Par.)* 248: 602–603.
14. Miller, O. J., Breg, W. R., Schmickel, R. D., and Tretter, W. 1961. Family with XXXXY male, leukemic male, and two 21-trisomic mongoloid females. *Lancet* 2: 78–79.

15. Miller R. W. 1963. Down's syndrome (mongolism), other congenital malformations and cancers among the sibs of leukemic children. *New Engl. J. Med.* 268: 393–401.
16. Nowell, P. C. and Hungerford, D. A. 1960. A minute chromosome in human chronic granulocytic leukemia. *Science* 132: 1947.
17. Penrose, L. S. 1934. A method of separating the relative aetiological effects of birth order and maternal age, with special reference to mongolian imbecility. Ann. Eugenics, 6: 108–127.
18. Penrose, L. S. 1938. A clinical and genetic study of 1280 cases of mental defect. *Spec. Rep. Ser. Med. Res. Coun.*, His Majesty's Stationery Office, London, No. 229.
19. Penrose, L. S. 1961. Mongolism. *Brit. M. Bull.* 17: 184–189.
20. Schull, W. J., and Neel, J. V., unpublished data.
21. Shapiro, S. 1962. A life table of pregnancy terminations and correlates of fetal loss. *Milbank Mem. Fund Quart.* 40: 7–45.
22. Shuttleworth, G. E. 1909. Mongolian imbecility. *Brit. Med. J.* 2: 661–665.
23. Steinberg, A. G. 1960. The genetics of acute leukemia in children. *Cancer* 13: 985–999.
24. Stewart, A., Webb, J., and Hewitt, D. 1958. Survey of childhood malignancies. *Brit. Med. J.* 1: 1495–1508.
25. Stewart, A. 1961. Aetiology of childhood malignancies: congenitally determined leukemias. *Brit. Med. J.* 1: 452–460
26. Stewart A. Personal communication.

PART V

Genetic Counseling

Some Aspects of Genetic Counseling

By Franz J. Kallmann, M.D., *Department of Medical Genetics of the New York State Psychiatric Institute, Columbia University, New York, N.Y.*

For the majority of therapeutically oriented physicians, genetic counseling holds no more attraction or significance than a number of other non-cure-directed activities that belong in the field of preventive medicine and are mainly designed to implement a public health program. A reflection of this low-priority rating may be seen in the limited degree of attention devoted to family guidance work in the medical school curriculum. Other indications are a certain vagueness in defining the genetic variety of counseling as "a specialized consultative procedure" in the management of families with some member presenting clinical symptoms of an apparently genetically determined deficiency or disease (Tips *et al.*, 1962), and the rather tacit assumption that this kind of counseling falls among the responsibilities of general family counselors.

As a corollary of this displacement phenomenon, considerable emphasis has been placed on the counselor's objective to help his charges "build a better world" (Michael and Meyerson, 1962). This compensatory trend has persisted in various disciplines, although family counselors might have learned from Allport's description of rationalism in professional guidance work that acceptance of reason in matters of opinion or conduct is reconcilable with the singular human ability to blend tentativeness with a firm commitment to chosen values (Allport, 1962).

In line with this relatively realistic principle and under the impact of recent advances in basic genetics, we may find it both timely and consistent with reason at this interdisciplinary conference to re-evaluate the extent to which current genetic counseling programs have become more knowledgeable, more resourceful, and perhaps even more personalized. As a geneticist who is interested in mental

health problems, may I add that the previously mentioned definition of genetic counseling calls for further qualification anyway.

Among the qualifying criteria is the axiom that in providing potentially stressful morbidity risk information for married couples and future parents, we have to give full consideration to the *emotional* state of the persons seeking guidance. By the same token, it is essential for counselors offering advice in regard to genetically rooted family problems that they be aware of their role in a *professional health service team* expected to be possessed of "a well developed sense of public responsibility" (Kallmann and Rainer, 1963).

In a great variety of counseling situations, this stipulation includes at the very least familiarity with the *simple facts of human genetics*—if family counselors or family doctors, who are most frequently consulted about common parenthood problems, do not want to be blamed for being a constant source of misinformation and mismanagement to their patients. In this conforming world of ours it is typical of people in general that they trust the word of authority, regardless of its degree of scientific validity. To the layman, any guidance counselor or psychotherapist fulfills the function of a physician and, therefore, belongs to those vaguely defined health education professions that represent the accepted authority on human genetics, whether he is adequately informed or not.

Apart from an understanding of special health risk statistics, a satisfactory working knowledge of *intrafamily dynamics* is needed in genetic counseling. One of the most complex tasks in this field of specialization is that of providing scientifically valid and psychologically innocuous guidance in problems of marriage and parenthood where there is a serious pathological condition in one or both of the families concerned. Undoubtedly, there are no more fateful decisions in a man's life than the one to marry and the one to assume parenthood. None is more consequential, either from the standpoint of prospective parents or their future children. It follows, therefore, that in the presence of uncertainty or doubt, whether well-founded or not, any family in an enlightened society should have access to competent advice.

In order to be of constructive help to unknowing or perplexed persons about to make a realistic choice in coping with their particular family problem, a qualified counselor is not only expected to elicit and discreetly evaluate the essential facts bearing on a family's decision, but also to attempt to understand the given persons' fears and hopes, defenses and rationalizations (Kallmann and Rainer, 1963). When confronted with a calculable health risk in a predictably unstable home, he will be guided by the general principle that, in this age of growing concern for the welfare of children, every child born "should be given a fair chance in life" (Slater, 1960).

It is understood, too, that the genetic section of the initial interview should include a *complete family history*, an essential part of any medical examination, although presently considered by some disciplines to be even more outdated than an exact psychiatric diagnosis. Whenever necessary, of course, the family history is to be augmented by such special cytogenetic techniques as chromosomal analyses or sex-chromatin counts.

As for the family inventory portion of the initial examination, one of the basic principles of conduct in genetic counseling is derived from the ancient medical maxim, *nil nocere*. Precisely because of the dire anxiety, guilt and potential unhappiness following upon only dimly understood verdicts regarding genetics, it is important to avoid giving information of this nature in an impersonal way—for instance, in the form of bare statistics or without empathic understanding of the motives and capacities of the persons who come for help. For this reason, such information should never be offered by mail or through third parties, but only after one or more extensive interviewing sessions.

Although these sessions need not be categorized as a formal program of psychotherapy, they will greatly gain in effectiveness if based on psychological understanding and conducted along established lines of psychiatric interviewing. In essence, sessions of this kind amount to *explanatory* or *manipulative* forms of short-term psychotherapy aimed at reducing anxiety and tension.

This cautious approach is justified by the common observation that even highly intelligent persons may regress to *immature levels* of emotion and thought, when frightened either by some factual morbidity risk information supposed to dispel their state of perplexity, or by the need to be evaluated in their roles as actual or prospective mates and potential parents. In such a frame of mind, people cannot be assumed to be capable of making important decisions in family matters without some reassuring measures of authoritative guidance and direction (Kallmann, 1953, 1956, 1958).

Another requirement of genetic counseling is that, when the stage is reached where a *program of action* can be formulated with reassurance, each step should be suited to the individual needs and resources of the given person or persons, both as to content and manner of presentation. In many instances, of course, the decisions to marry and to have children will have to be dealt with as two separate problems. If one or the other question raises serious doubts, it will be advisable to be guided by the truism that healthy children (healthy in terms of genetic endowment, developmental circumstances and emotional climate in the home) are both the product and the source of health in the parents.

Further responsibilities of a genetic counselor include familiarity with *current adoption procedures*. When preference for a couple's

childlessness is indicated on genetic grounds, there emerges the need for a frank discussion of the modes of operation and potential risks of these procedures. Certainly, the hazards of unwittingly adopting the natural child of two schizophrenics overshadow the risk of harelip or diabetes in a child of one's own. Justifiable fears of this kind cannot be allayed by the glib assertion that the effect of heredity in man is negligible or will forever remain beyond our control (Kallmann, 1962; Kallmann, 1963).

How important it is for a genetic counselor to use *procedural flexibility, supportive guidance* and a *realistic attitude* in his work may be illustrated by the following examples:

1. In the first case, an emotionally bewildered, socially prominent couple, an artist married to a professional man, was referred by a private psychiatrist for genetic counseling inasmuch as they were known to be *double first cousins*. They had put off having children for 9 years because of vaguely formulated fears engendered by the family physician and the "older folks" in their family. Fortifying their anxiety with many other rationalizations, they came to believe that there was no room in their lives for children. However, their real wishes, which were slowly worked out with them, were just the opposite. Study of the family backgrounds and sympathetic reassurance gave the couple both the will and the courage to have a child. Born in 1961, the baby proved to be normal and healthy (fig. 1), thus conferring a full measure of happiness on all concerned.

2. In another case (fig. 2), an intelligent and attractive young lady, the only *hearing member of a deaf family*, came to see us because she wanted to get married. Her main problem was that she felt obliged to marry within her highly inbred religious community. This emotionally overvalued notion seemed to stem from her belief that she was the least fortunate and most poorly adjusted of the 3 children of her deaf parents, presumably because she had been utterly rejected by all the deaf members of the family. Apparently, some relatives and neighbors had suspected her deaf but beautiful mother of having solved her genetic problem, after giving birth to 2 deaf children, with the help of what Fraser (1963) called an NID (to distinguish a natural insemination donor from an AID or artificial insemination donor). At any rate, we advised the young lady in question to marry a *complete outsider*.

Several years later, she returned with her husband for parenthood counseling, requesting advice regarding the risk of deafness if they were to have a child. Here statistics were not very helpful. At any rate, the woman's background, specific fears and social conflicts had to be considered along with the genetic probabilities. In view of the fact that her dread of deafness outweighed every other consideration, she was encouraged to follow her plan to *adopt a child* rather than bear one.

FIGURE 1.—Healthy Child of Double First Cousins

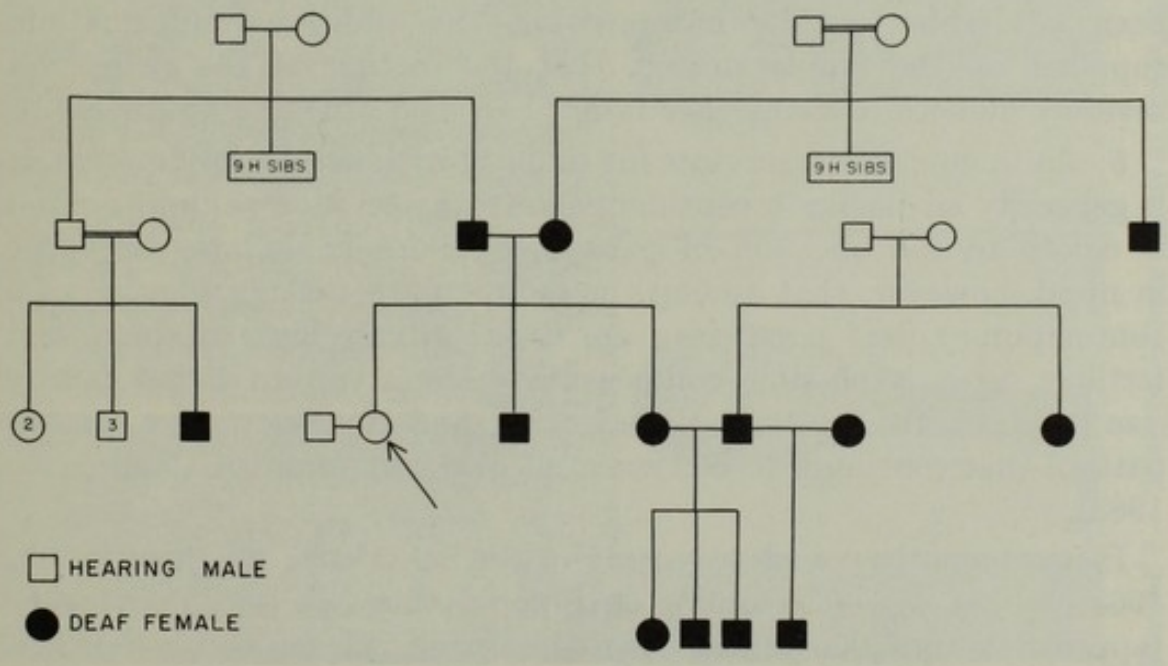


FIGURE 2.—Counseling Problem in the Hearing Member of a Deaf Family

3. A different genetic counseling picture was presented by the mother of a pair of mentally retarded twins of the 1-egg variety (fig. 3), who were congenitally blind due to *microphthalmia* and had, in addition to 4 normal brothers and sisters, 2 older sibs with equally

undeveloped eyes (Kallmann, 1953). The twins were born at full term and without instruments, although the mother was already 43 years old at the time. Both parents had normal vision, but they were children of first cousins.

Following the twins' birth, the mother insisted upon a *sterilizing operation*—with our unqualified recommendation. Her earlier requests for sterilization had been rejected by 2 other medical centers on the theory that the defects of her first 2 eyeless children might have been "an unfortunate coincidence." Obviously, this explanation ignored the principles of *recessive inheritance* as well as the possibilities of a *chromosomal abnormality*, perhaps in the D group of chromosomes 13–15.

4. The need for adequate clinical histories in the teamwork of genetic counselors may be illustrated by another pair of blind and grossly retarded twins (fig. 4), who come from an average middle-class family and are of the 2-egg type. Born prematurely in 1952, with birth weights of 1,230 and 1,440 gm., they are low-grade defectives with IQ's below 20 and completely concordant for the cicatricial form of *retrolental fibroplasia*. Upon delivery both twins were placed in incubators and given oxygen for 6 to 8 weeks.

With the *exogenous* origin of this perinatal condition fully established (Kinsey, 1956), the mother's request for sterilization following the birth of such a defective pair of twins would not under ordinary circumstances be warranted. In this particular case, it might have been advisable to take into account the additional information, supplied by the family doctor, that the mother of the twins was severely diabetic and that her father had died of the same disease.

5. As to counseling services for families with *early total deafness*, it is generally advisable to consider each family on its own merits when it comes to the question of parenthood. It should also be borne in mind, however, that among the factors perpetuating the size of a contemporary deaf population are the relatively high marriage and fertility rates, a possible compensatory trend toward larger family size in units with at least 1 deaf child, and the assortative mating pattern that continues to be typical of deaf communities (Kallmann, 1963).

In our recently completed study of the genetics of deafness (Sank, 1963), about one-half of the deaf population of New York State (approximately 6,000 persons over the age of 12) had to be classified as sporadic cases that seemed to be due at least in part to exogenous causes, to complicated modes of inheritance, or to new genetic mutations. With dominant genes accounting for only about 10 percent of all forms of deafness, most inherited deafness cases were assumed to follow the autosomal-recessive mode of inheritance. Of particular interest for counseling purposes was the observation of a *more variable*

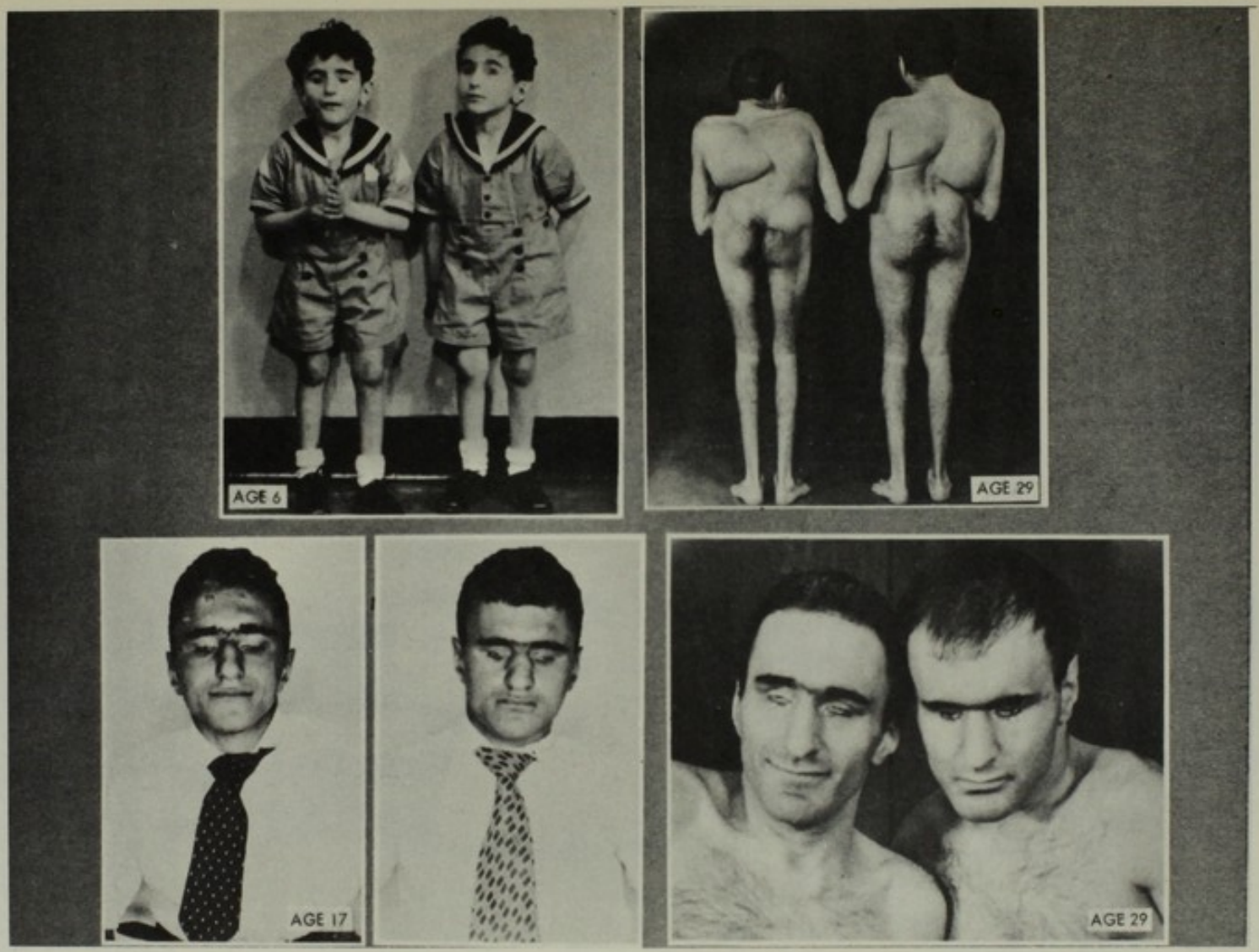


FIGURE 3.—Microphthalmic Twins at the Ages of 6, 17 and 29 Years (Mansfield State Training School)

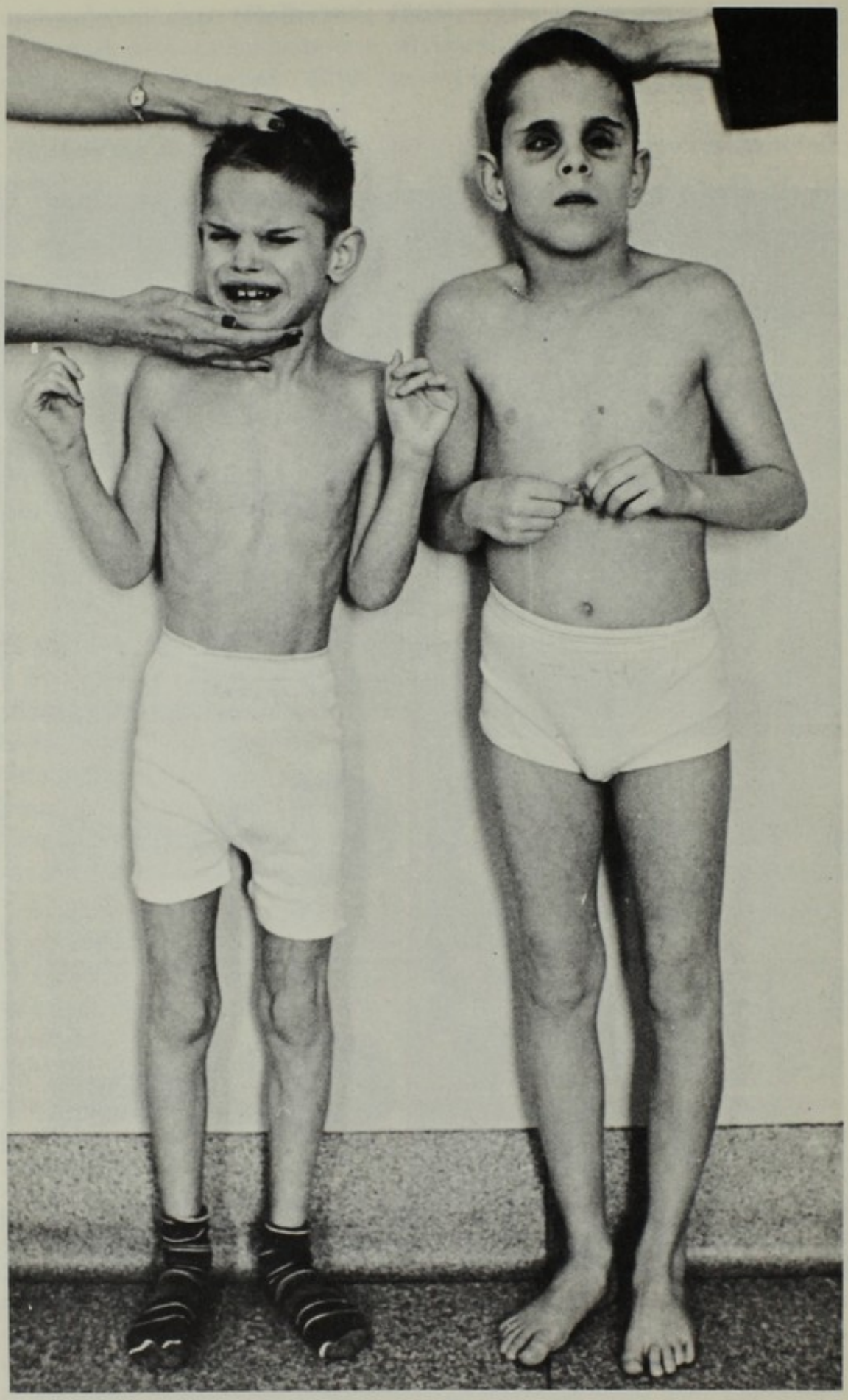


FIGURE 4.—Concordance for Retrolental Fibroplasia and Mental Deficiency in the E. Twins (I.Q.'s below 20)

expressivity pattern in the dominant than in the recessive forms of deafness.

The pedigree of one of our twin families (fig. 5) may serve to illustrate variations in the expression of *dominant* genes, including hearing losses which are either of limited extent or expressed only unilaterally or later than in early life. This family contains a dual mating unit with five children, among them a pair of 1-egg twins concordant for partial ocular albinism and clinically discordant for early total deafness. Of the 3 sibs of these twins, the oldest sister is deaf and heterochromic, like the mother and some members of her family. The two younger sibs have normal hearing, although the girl has partial ocular albinism, and the boy had one eye enucleated because of a histopathologically verified retinoblastoma (Sank, 1962).

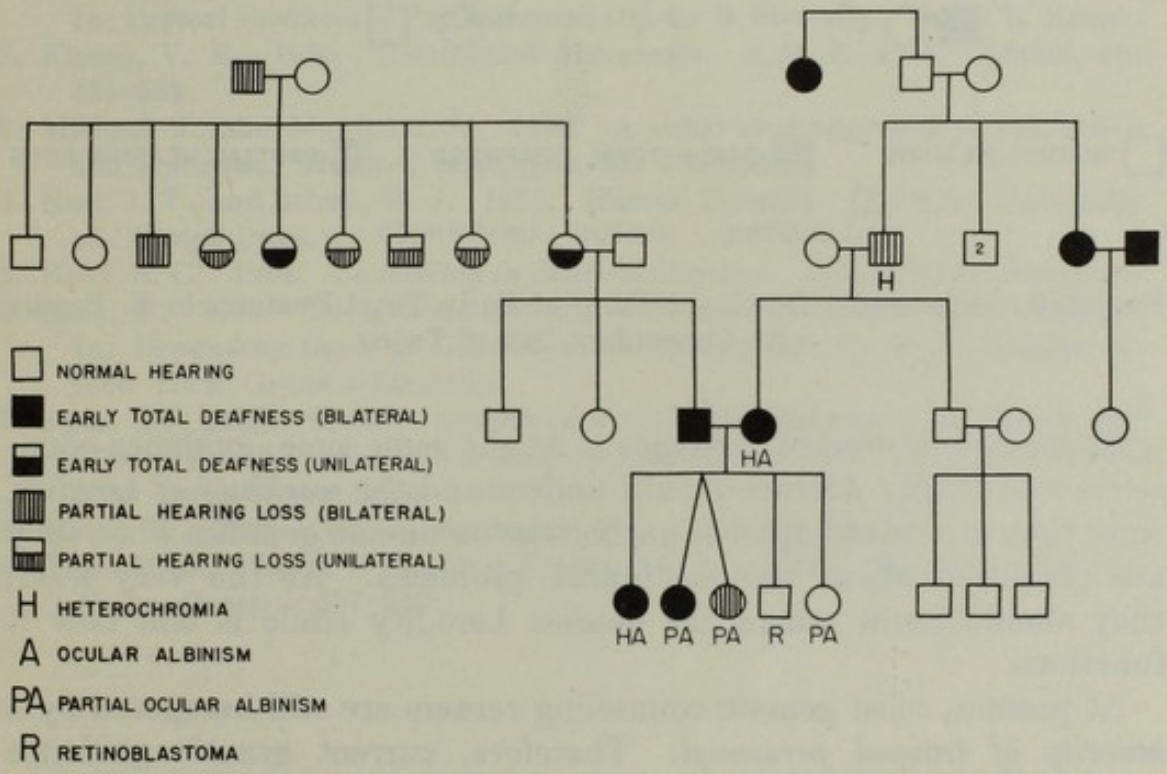


FIGURE 5.—Apparently Dominant Form of Early Total Deafness Showing Incomplete Expressivity and Association with Ocular Albinism, Heterochromia and Retinoblastoma in the R. Family

6. In another family which included a concordant pair of opposite-sex twins (fig. 6), 9 cases of *deafness* occurred in consanguineously related persons and were fully expressed, except for 1 case of partial hearing loss in a paternal aunt. Of the 3 dual matings in this family, one produced deaf offspring only, one hearing offspring only, and one a combination of deaf and hearing children—thus providing us with a striking example of the *technical complexities* encountered in genetic counseling.

In conclusion, it may again be stressed that, as a member of a health service team, a professional geneticist is called upon to *empathize*

735-712 O—65——26

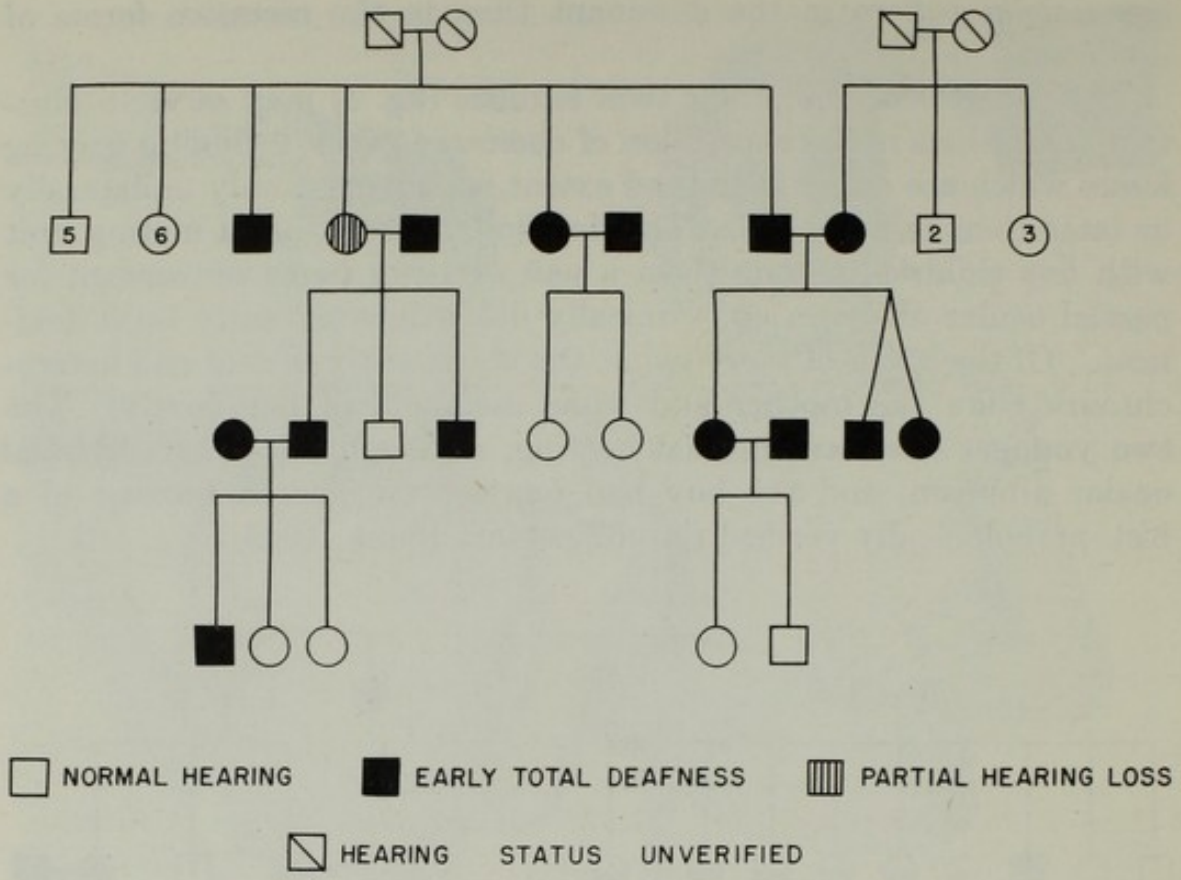


FIGURE 6.—Apparently Dominant Form of Early Total Deafness in E. Family with Concordant Pair of Twins

with persons in need of guidance. At the same time, guidance counselors and family doctors should understand the *workings of heredity*, since they are looked upon as authorities on human genetics when they are consulted about various family problems. At the very least, they should know where the nearest heredity clinic is and how it functions.

At present, most genetic counseling centers are still hampered by a *scarcity of trained personnel*. Therefore, current genetic guidance programs cannot really be expected to be free of *various imperfections*. There can be no question, however, that the constructive management of genetic family problems requires either *geneticists* who are experienced in counseling techniques or *family guidance workers* who have had adequate training in genetics.

WELLCOME LIBRARY INSTITUTE

REFERENCES

1. Allport, G. W. 1962. Psychological models for guidance. *Harvard Educ. Rev.* 32: 373–381.
2. Fraser, F. C. 1963. On being a medical geneticist. *Am. J. Human Genet.* 15: 1–10.
3. Hammons, H. G. (ed.). 1959. *Heredity Counseling.* New York: Paul B. Hoeber.

4. Herndon, C. N. 1955. Heredity counseling. *Eugen. Quart.* 2: 83–89.
5. Kallmann, F. J. 1953. *Heredity in Health and Mental Disorder.* New York: W. W. Norton.
6. Kallmann, F. J. 1956. Psychiatric aspects of genetic counseling. *Am. J. Human Genet.* 8: 97–101.
7. Kallmann, F. J. 1958. Types of advice given by heredity counselors. *Eugen. Quart.* 5: 48–50.
8. Kallmann, F. J. 1961. New goals and perspectives in human genetics. *Acta Genet. Med. et Gemell.* 10: 377–388.
9. Kallmann, F. J. 1962. Genetic research and counseling in the mental health field, present and future. In: *Expanding Goals of Genetics in Psychiatry* (ed. by F. J. Kallmann). New York: Grune & Stratton.
10. Kallmann, F. J. 1963. Main findings and some projections. In: *Family and Mental Health Problems in a Deaf Population* (ed. by J. D. Rainer et al.). New York: Comet Press.
11. Kallmann, F. J., and Rainer, J. D. 1963. Psychotherapeutically oriented counseling techniques in the setting of a medical genetics department. In: *Topical Problems of Psychotherapy* (ed. by B. Stokvis). Basel: S. Karger.
12. Kinsey, V. E. 1956. Retrolental fibroplasia. *A.M.A. Arch. Ophthal.* 56: 481–543.
13. Michael, J., and Meyerson, L. 1962. A behavioral approach to counseling and guidance. *Harvard Educ. Rev.* 32: 382–402.
14. Neel, J. V., and Schull, W. J. 1954. *Human Heredity.* Chicago: University of Chicago Press.
15. Reed, S. C. 1955. *Counseling in Medical Genetics.* Philadelphia: Saunders.
16. Sank, D. 1962. The genetic and adjustive aspects of early total deafness. In: *Expanding Goals of Genetics in Psychiatry* (ed. by F. J. Kallmann). New York: Grune & Stratton.
17. Sank, D. 1963. Genetic aspects of early total deafness. In: *Family and Mental Health Problems in a Deaf Population* (ed. by J. D. Rainer *et al.*). New York: Comet Press.
18. Slater, E. 1960. Galton's heritage. *Eugen. Rev.* 52: 91–103.
19. Tips, R. L., Meyer, D. L., and Perkins, A. L. 1962. Genetic counseling. *Eugen. Quart.* 9: 237–240.

U.S. GOVERNMENT PRINTING OFFICE: 1965 O—735–712

THE ROYAL SOCIETY

FOR THE PROMOTION

OF HEALTH

Founded 1876

LIBRARY REGULATIONS

(a) Books, periodicals and pamphlets may be borrowed by Fellows, Members, Licentiate Members, Associates and Affiliates personally or by a messenger producing a written order. The person to whom such publications are delivered shall sign a receipt for them in a book provided for that purpose.

(b) Publications may be borrowed through the post upon a written order. An acknowledgement of the receipt of such publications must be made on the form provided and returned to the Society immediately. Failing this, it will be assumed that the borrower has received the books, for which he will accordingly be held responsible. The postage of publications returned to the Society shall be defrayed by the borrower.

(c) A borrower may not have more than three publications in his possession at one time.

(d) A borrower will be considered liable for the value of any publication lost or damaged while on loan to him, and, if it be a single volume or part of a set, for the value of the whole work thereby rendered imperfect. Marking or writing in the publications is not permitted, and borrowers are requested to call attention to damage of this character.

(e) **Books and pamphlets may be retained for twenty-eight days. Periodicals may be retained for fourteen days. Applications for extension of the loan period must be made in writing before its expiry.** This loan period cannot be extended without application being made to the Secretary of the Society whose decision in this matter must be accepted as final and binding.

(f) Books and pamphlets added to the Library will not be lent until after the expiry of one month from the date received. The current number of a periodical may not be borrowed.

(g) Borrowers retaining publications longer than the time specified, and neglecting to return them when demanded, forfeit the right to borrow until they be returned, and for such further time as may be ordered by the Council.

Any borrower failing to comply with a request for the return of a publication shall be considered liable for the cost of replacing it, and the Council, may, after giving due notice to him, order it to be replaced at his expense.

No publication may be reissued to the same borrower until at least seven days have elapsed after its return, neither may it be transferred by one borrower to another.

(h) **Publications may not be taken or sent out of the United Kingdom.**

(i) Publications returned through the post must be securely packed and adequately protected.

(j) Parcels should be addressed: THE LIBRARIAN,
THE ROYAL SOCIETY OF HEALTH
~~90 BUCKINGHAM PALACE ROAD~~, LONDON, S.W.1 X 7EN.
O.B., Truro. 13 GROSVENOR PLACE,

—January, 1965