AN EPIDEMIOLOGIC APPROACH TO

Inherited Disease Susceptibility

THOMAS D. DUBLIN, M.D., Dr.P.H., and BARUCH S. BLUMBERG, M.D., Ph.D.

THE SCIENCE of epidemiology has been firmly established on the concept of multiple causation of disease. Each of three major etiological factors—a causative agent, environmental conditions, and host susceptibility—plays a determinative role in disease occurrence. Traditionally, however, the epidemiologist, preoccupied with infectious processes, has concentrated on the study of the first two factors: causative agents and environmental conditions. As a rule, he has devoted far less attention to the peculiarities of the individual host or of the population in which a disease occurs, or fails to occur.

To be sure the epidemiologist is accustomed to studying the distribution of disease according to age, sex, race, occupation, ethnic background, and other attributes of those affected; one of the primary purposes of such analyses being the development of etiological hypotheses concerning the disease. These analyses have been indispensable in identifying exposure to a specific environmental hazard or in understanding the immunological implications of specific acquired immunity, particularly in those conditions in which infection and disease do not coincide. Little is yet known, however, of the details of natural or innate host resistance or of the factors which affect host susceptibility and are characteristic of the individual or population prior to exposure to infection.

Epidemiologists are now concerned with a broad array of diseases in which a specific invading parasite or toxic substance cannot be identified as of etiological significance. In such conditions innate characteristics of the host assume critical importance. In some, long periods of latent abnormal metabolism or even frank tissue pathology precede manifest disease. Diabetes, rheumatoid arthritis, atherosclerosis, ischemic heart disease, and allergic phenomena, as well as many other diseases, fall into this category. In studying these conditions the epidemiologist is challenged by the tasks of identifying the susceptible individual, determining the basic defect, and then assessing the factors or stresses which break through the reserve of the susceptible individual and cause disease. To achieve this goal it is necessary to seek new clues to disease susceptibility and to find new ways to identify, prior to the appearance of manifest disease, those individuals who are subject to an increased risk.

From a different vantage point, the medical geneticist has recently begun to approach the same or kindred problems. He has found that genes may influence the occurrence of disease in a variety of ways. Some diseases are under direct genetic control. That is, if a given gene (or genes) is present, then, under most environmental conditions, the disease will be manifest. Such diseases are designated as dominant, recessive, co-dominant, and so forth. The significance of these terms is discussed elsewhere (*1*). Most of the diseases transmitted in this manner are rare (although under certain circumstances they may occur in high frequencies in a population), and their inheritance usually follows Mendelian segregation. They include such entities as inherited deaf mutism, galactosemia, phenylketonuria, various skeletal deformities, and many others. The study of these

Dr. Dublin is chief of the Epidemiology and Biometry Branch, and Dr. Blumberg is chief of the Section on Geographic Medicine and Genetics, National Institute of Arthritis and Metabolic Diseases, Public Health Service. *This paper was presented at the 88th annual meeting of the American Public Health Association in San Francisco, October 31, 1960.*

conditions has been the main concern of medical genetics for many years (*2*).

For many common diseases in which the evidence for inherited predisposition is impressive the genetic mechanisms involved appear to be more complex. It has been postulated that tuberculosis, diabetes, rheumatoid arthritis, and hypertension, for example, depend on "polygenic" factors and like such "normal" traits as height, weight, skin color, and intelligence are due to the interaction of several and perhaps many genes.

A characteristic of polygenic traits is the considerable effect of environment superimposed on the inherited component. These traits are often continuous (or metric) in character with the values occurring according to a Gaussian distribution. This is in contrast to the discontinuous, or discrete, distribution of the rare genetic traits mentioned above where two or more classes can be readily identified.

It is not unlikely that some or all of the genes contributing to the polygenic inheritance of a disease act so as to influence susceptibility or protection against the diseases. Thus individuals with the susceptible genotype or genotypes will not be affected unless exposed to the causative agent under environmental conditions which precipitate the illness. Or, on the other hand, when exposed to a hazard those in the population who are genotypically susceptible will be more likely to contract the disease while those with the alternate genotypes will be relatively more resistant.

Needed for a fuller appreciation of inherited susceptibility is a better understanding of the expression of the individual genes which make up the polygenic system. Some information on these genes and the traits they determine may be obtained from the study of the genetic polymorphisms.

Ford (*3*) has defined a polymorphism as the occurrence together in the same habitat of two or more inherited discontinuous forms of a species in such numbers that the form in least frequency could not be maintained by recurrent mutation alone. The use of a familiar example may help to explain this definition (*4*). There is a gene common in some African and other populations which acts on the beta polypeptide chain of hemoglobin to substitute the amino acid valine for the glutamic acid found in the beta chain of normal human hemoglobin. This alteration causes the formation of sickle-shaped red blood cells under certain conditions. When an individual is homozygous for this gene nearly all of his hemoglobin is of the abnormal sickling type. Heterozygotes have some sickling and some normal hemoglobin. In African and other populations where this gene is common, there are three easily distinguishable (discontinuous) forms of humans:

1. Those who have all normal hemoglobin, genotype Hb^A/Hb^A.
2. Those who have sickle-cell hemoglobin and who suffer from sickle-cell disease, genotype Hb^S/Hb^S.
3. Those who have some sickle-cell hemoglobin and some normal hemoglobin but who have no apparent disease, genotype Hb^S/Hb^A.

Why are there three phenotypes and two allelic genes for the hemoglobin trait? Would not one have been sufficient? The explanation appears to be that these types arose as a result of differences in susceptibility to an important selective factor in the tropical environment, namely, malaria. This conclusion has been derived from the following line of reasoning.

Sickle-cell disease is characterized by an acute anemia which is often fatal in childhood. It has been estimated that the viability of children with the sickle-cell disease (Hb^S/Hb^S) is one-fifth of that of children with the Hb^A/Hb^A genotype. Every time an Hb^S homozygote dies, two genes are removed from the population. In a tribe where the sickle-cell trait frequency approaches 40 percent, about 4 percent of children are born with sickle-cell disease Hence, about 15 percent of the sickle-cell genes would be eliminated in each generation, and it would not require many generations before the frequency of the Hb^S gene would have decreased to a level where it was maintained by recurrent mutation alone. On the basis of known mutation rates in man, it is very unlikely that the present levels of the Hb^S gene in these African populations could be maintained by recurrent mutation alone. It has been suggested that the reason for the maintenance of these high frequencies is that the

heterozygous child has a greater relative protection against falciparum malaria than does the homozygote with normal hemoglobin. Hence, the sickling gene is characteristic of this population and has become so because of the selective effect of the mass killing disease malaria over the course of many generations.

As long as protection against malaria (and/or other selective factors not yet recognized) offsets the losses due to death of individuals homozygous for sickling, the polymorphism will remain balanced. In 1930, Fisher in "The Genetical Theory of Natural Selection" (*5*) showed that conditions resulting in an advantage to the heterozygote could maintain a balanced polymorphism in the population even though the relative advantage to the heterozygote was not very great. There are situations where the polymorphism is not balanced but transient, with gene frequencies changing markedly from generation to generation. This appears to have occurred when populations with the sickle-cell polymorphism were transported to areas where falciparum malaria was absent. In the 17th century, most of the slaves transported to the Dutch New World possessions of Curacao and Surinam (Dutch Guiana) appeared to have been shipped from the same location on the Gold Coast. The island of Curacao had no malaria, while the disease was very common on the mainland colony of Surinam. When allowances are made for different degrees of gene mixture, there appears to be considerably less of the sickle-cell trait in the Curacao population of African descent than in the equivalent Surinamers. We may infer from this that the frequency of the gene has decreased in Curacao where the selective influence of malaria has been removed, while it is still high in Surinam where falciparum is endemic (*6*). During the period when the frequency of the sickle-cell gene was decreasing, the polymorphism was transient. There may be many such examples in human populations, but it is, of course, difficult to detect them when observations are made at one point in time.

From an epidemiologic viewpoint this example of a balanced polymorphism illustrates how a genetically determined trait and a disease associated with this trait, sickle-cell anemia, are maintained in a population and are dependent on a relatively delicate and complex balance of genetic and environmental factors. Sickle-cell anemia is maintained because in single dose the gene determining this condition provides protection against a major killing disease. If there were no selective advantage to the heterozygote, sickle-cell anemia would be a rare disease in West Africa with its frequency ultimately determined by recurrent mutation alone. Other forms of heterozygote advantage, such as differential fertility, may also play a role in the maintenance of the polymorphism.

In recent years a significant number of presumptive polymorphisms have been identified in human populations. For the most part such traits have aroused greater interest among anthropologists than among epidomiologists. They have aided the anthropologists in characterizing populations of divergent or similar genetic background but only rarely have efforts been made to determine whether such polymorphisms can be associated with important disease problems within those populations. The identification of disease-polymorphism associations demands the type of extensive controlled observations of populations that have been the concern of epidemiologists for generations.

With respect to malaria at least two other polymorphisms have been recognized which may bear on malaria susceptibility. Glucose-6-phosphate dehydrogenase (G6PD) deficiency appears to confer protection against falciparum malaria (*7,8*). However, individuals with this trait develop a sometimes fatal acute anemia when they are exposed to the fava bean and flower, although under other circumstances they are apparently normal. They also develop hemolytic crises after the ingestion of a large variety of drugs, including primaquine. It is, of course, unlikely that the acute anemia resulting from the ingestion of these drugs could be a cause of natural selection. However, it is not improbable that there are natural products which are ingested that may have a similar effect and could affect natural selection. The fava bean example is one of these. Another polymorphism which may also act on malaria is thalassemia which occurs in some Mediterranean populations. Although at present there is no conclusive evidence that there is a protective

influence of this trait, material collected in Sardinia is consistent with this surmise (*9*). Hence, in malaria at least two and perhaps three genes are known which have some influence on protection, and the nature of the genes making up the polygenic generalized inheritance is being identified.

The polymorphisms currently present in human populations could be due to a variety of causes. Some may have been generated and are being maintained by selective forces which affect the differential reproductivity of a population, that is, by widespread diseases of public health importance. The term "differential reproductivity" is used here in the broad sense and includes all factors which permit one segment of the population to contribute reproducing members to the next generation to a different degree than other segments of the population. This will lead to changes in gene frequency when the segments are characterized by different genotypes.

Some polymorphisms may have been generated by diseases which were important in the past but which are no longer of selective value. The polymorphism will remain in a population dependent on generation time and the remaining selective values of the phenotypes. This, incidentally, points up the importance of studying primitive societies. The selective agent might still be obvious in these populations although it is not evident in more highly developed societies. It also points out a very interesting branch of medical history: the study of diseases in primitive populations prior to or at the time of the arrival of Western chroniclers. There is excellent evidence that diseases have changed radically over the course of generations. This is probably due to evolutionary changes in the causative agent and in the host resistance of the human population.

It is also possible that "neutral genes" exist, that is, genes which have no important physiological effect and whose frequencies are determined by chance. Neutral genes involve the concept of "genetic drift," but the discussion of this mechanism is beyond the scope of this paper. In some cases selective advantages giving rise to a polymorphism may be very slight indeed and difficult to detect. When such a situation prevails, it would be necessary to study the genotypes of a very large number of individuals with and without disease in order to establish a relationship.

How common are these polymorphisms? Some human biochemical traits that fall within this definition are listed on p. 504. The trait is included if the phenotype in lowest frequency is found in approximately 1 percent or more of some populations and only those cases in which the genetics is fairly well determined have been included. Some of the genetic systems determining relatively rare recessive inherited disease (such as cystic fibrosis) would also fall into this definition but are not included in the list. There are undoubtedly many such polymorphisms; several are discovered each year.

Polymorphic systems which are convenient for study have the following characteristics:

1. The phenotypes are easily distinguished and preferably are qualitatively different.

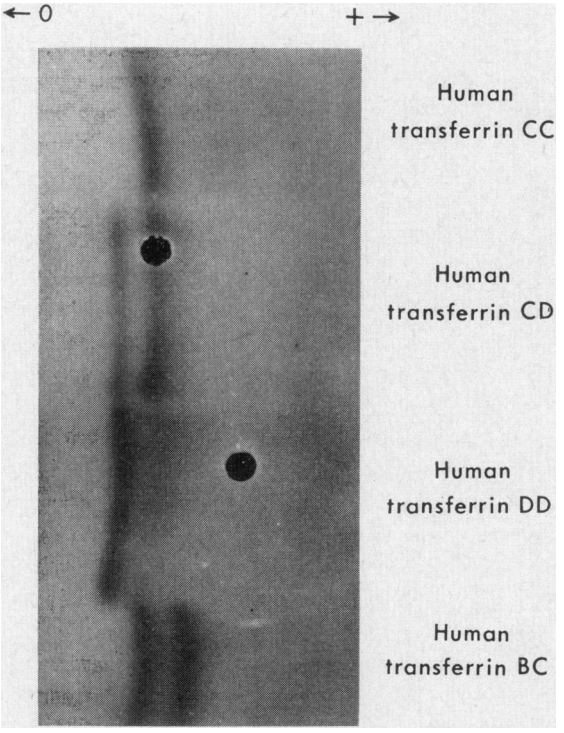

Figure 1. Human serum transferrins. The photograph of a section of a starch gel electrophoretic run shows differences in mobility of four transferrin phenotypes. The heavy black dots are index marks; the upper dot indicates the position of transferrin C. The origin is to the left and the positive pole to the right.

2. The expression of the gene is more or less invariable during life (or at least after the perinatal period). An individual's phenotype can be determined under all or most environmental conditions.

3. The genetics of the system is explicit. The genotype or genotypes can be determined directly from the phenotype.

4. The genes are expressed as chemical units, such as enzymes, antigens, antibodies, co-factors, proteins, and mucopolysaccharides, which can be isolated for characterization and biochemical and physiological studies. Obvious relations to a physiological, pathological, or selective factor are of particular value.

5. The phenotypes are fairly common in the population.

Two examples of such traits are shown in figures 1 and 2. The transferrins are a family of serum proteins which bind iron (*10*). Transferrins of different electrophoretic mobility can be distinguished by means of starch gel electrophoresis (fig. 1). Most Europeans have a single protein termed "transferrin C." Many Africans, some American Negroes, and members of other populations have in addition a slow-moving band termed "transferrin D." Some Europeans, Americans, and Asiatic people have a fast-moving band termed "transferrin B." These proteins are determined by three genes: Tf^{C}, Tf^{D}, *and* Tf^{B} at the same locus, and are inherited as simple Mendelian traits fully expressed when the gene is present in a single dose. Biochemical studies on these proteins are in progress (*11*).

A "new" polymorphism which also illustrates some of the characteristics detailed above has been identified recently (*12*). The serum of a patient who had received many blood transfusions was found to contain a material with some of the characteristics of an antibody. It reacted with an alpha globulin "antigen" in the serum of some but not all normal subjects. The presence or absence of the "antigen" is under direct genetic control. Individuals homozygous for a gene *Ag* (genotype *Ag/Ag*) do not react; individuals homozygous for the dominant allele Ag^{A} or heterozygotes (genotypes Ag^{A}/Ag^{A}, Ag^{A}/Ag) do react. The reactions with positive and negative serums in an agar gel Ouchterlony plate are shown in figure 2.

Are all these polymorphisms simple curiosities or are they of interest to the physician and epidemiologist? The answer probably is that some are and some aren't, and it is difficult to decide which to study. With one or two notable exceptions the presently known polymorphisms have been identified as the result of chance observation rather than purposeful search. Intensive explorations of these polymorphisms offer opportunities for the epidemiologist, geneticist, biochemist, immunologist, and anthropologist to join forces with at least some promise of gaining better insight into the interrelated roles of genetic and environmental factors in disease causation and disease susceptibility.

Associations between the blood groups and various disease entities have been studied in re-

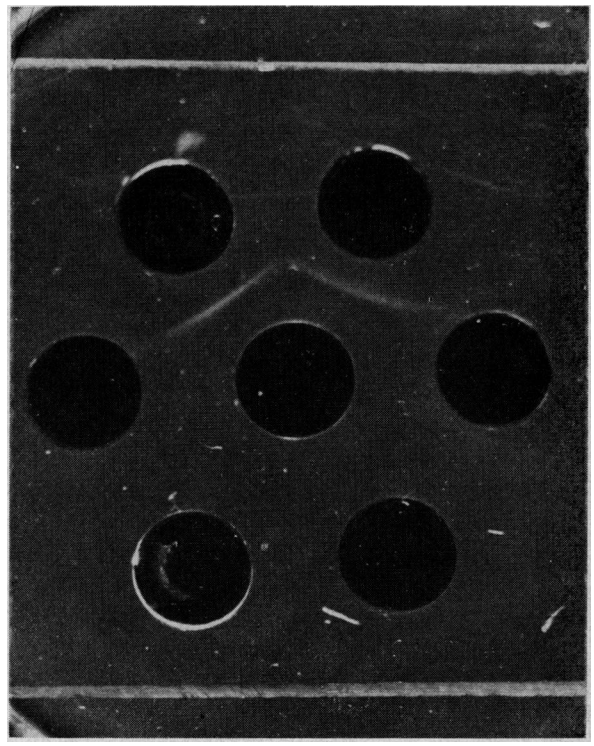

Figure 2. Precipitin reaction in agar gel Ouchterlony plate between an "antibody" present in the serum of a patient who had received multiple transfusions and an alpha globulin "antigen" in the serum of some but not all normal humans. The serum containing "antibody" is in the center well and the serum containing "antigen" is in the peripheral wells. There is a precipitin reaction between the "antibody" and the serums in the top two wells but not in the others.

cent years and have been summarized by Roberts (*13*). In some of these selection is of a small order and other forces may also have operated to maintain the blood group polymorphisms. Duodenal ulcer occurs as high as 40 percent more frequently among individuals with blood group O than in those belonging to the other ABO groups. Interesting though less marked associations between blood group A and gastric cancer, pernicious anemia, and diabetes mellitus have also been observed. Using the peptic ulcer association as a model, one need not assume, however, the blood group O itself enters into the susceptibility process any more than ABH bestows protection. (It might be added parenthetically that this possibility has not been extensively studied.) A gene may have multiple effects, and its most easily detected expression may not be the same as its most important physiological role.

Recently, the National Institute of Arthritis and Metabolic Diseases, Public Health Service, has begun to explore some of the scientific leads briefly summarized here. The newly established Epidemiology and Biometry Branch has incorporated as an integral component a Section on Geographic Medicine and Genetics, and the attention of the staff will be focused on field investigations coupled with appropriate laboratory and clinical studies in this area.

As examples of work underway or in process of development we have selected several studies for brief comment. Glynn, Glynn, and Holborow (*14*) have found that there are higher frequencies of salivary nonsecretors of the ABH substances in a population of persons with rheumatic fever than in a control group. From the data of these workers and that of Clark and associates (*15*), there also appears to be a lower frequency of individuals of blood group O in the rheumatic fever population. There is also a relationship with the Lewis substance, but the genetics of this system is still not completely elucidated. We are attempting to confirm these studies and extend the observations to other populations. If confirmed it would appear that at least two polymorphic systems are related to susceptibility to rheumatic fever. It is, of course, well known that beta hemolytic streptococcus infection may precipitate rheumatic fever attacks, but it is still not known why some individuals who are so infected succumb to this disease while others do not.

On the basis of chemical and clinical studies there is evidence that the inherited ability to taste phenylthiocarbamide and related compounds may be associated with susceptibility to certain types of thyroid disease (*16*). We have undertaken exploratory studies and are hoping to initiate a field investigation in one of the high-incidence goiter regions of the world, possibly in the equatorial Andes, to test this hypothesis by epidemiologic studies. The roles played by genetic susceptibility and by naturally occurring environmental goitrogens,

Some known polymorphisms and disease states or physiological systems with which they are thought to be associated

Polymorphism	*Disease or physiological state*
Blood groups:	
A_1A_2BO	Peptic ulcer, cancer of stomach, pernicious anemia, diabetes, ?rheumatic fever, ABO incompatibility, ?fertility.
Rh	Erythroblastosis fetalis.
MNSs	?Fertility.
Kidd	?
Kell	?
Duffy	?
Lewis	?Rheumatic fever.
Lutheran	?
Diego	?
Sutter (Js)	?
P	?
Hunter and Henshaw	?
Secretor state	?Rheumatic fever.
Hemoglobin S	Anemia, malaria.
Hemoglobin C	Anemia, ?malaria.
Glucose-6-phosphate dehydrogenase (G6PD).	Favism, drug intoxication, malaria.
Thalassemia	Anemia, ?malaria.
Haptoglobins	Hemoglobin metabolism, ?infection.
Transferrins	?Iron binding, ?infection.
Gamma globulins.	
Alpha globulin antigen Ag-(a+).	?Transfusion reaction.
Beta-amino-isobutyric excretion (BAIB).	?Leukemia, ?radiation.
Taste of phenylthiocarbamide (PTC).	?Thyroid disease.
Serum cholinesterase	Drug sensitivity.
Isoniazid inactivation	Drug sensitivity.

as well as by iodine deficiency, in the etiology of endemic goiter, a disease estimated to affect 200 million persons, have not yet been fully explored and offer promising opportunities for epidemiologic research (*17*,*18*).

In summary, what then is the significance of inherited susceptibility to epidemiologic research? It can be seen from the foregoing that the dynamics of this characteristic has not yet been fully elucidated, but progress in this direction is accelerating rapidly. Inherited disease susceptibility is probably due to many genes acting in concert, and some of these genes may form polymorphic systems. For some diseases we are approaching an understanding of at least some of the processes; for others the prospects of gaining better insight into the mechanism accounting for innate resistance or susceptibility are promising. The epidemiologist could make a significant contribution to studies in this area and his understanding of disease occurrence and distribution would be enhanced by a fuller appreciation of present-day medical genetics.

REFERENCES

(*1*) Allison, A. C., and Blumberg, B. S.: Dominance and recessivity in medical genetics. Am. J. Med. 25: 933–941, December 1958.

(*2*) Stern, C.: Principles of human genetics. Ed. 2. San Francisco, W. H. Freeman & Co., 1960.

(*3*) Ford, E. B.: Genetics for medical students. London, Methuen & Co., Ltd., 1942.

(*4*) Allison, A. C.: Malaria in carriers of the sickle-cell trait and in newborn children. Exper. Parasitol. 6: 418–447, July 1957.

(*5*) Fisher, R. A.: The genetical theory of natural selection. Ed. 2. New York, Dover Publications, Inc., 1958.

(*6*) Jonxis, J. H. P.: The frequency of haemoglobin S and haemoglobin C carriers in Curacao and Surinam. *In* Abnormal haemoglobins. Oxford, England, Blackwell, Basil & Mott, Ltd., 1959, pp. 300–306.

(*7*) Allison, A. C.: Glucose-6-phosphate dehydrogenase deficiency in red blood cells of East Africans. Nature 186: 531–532, May 14, 1960.

(*8*) Motulsky, A. G.: Population genetics of glucose-6-phosphate dehydrogenase deficiency of the red cell. *In* Proceedings of the Conference on Genetic Polymorphisms and Geographic Variations in Disease. In press.

(*9*) Carcassi, U., Ceppellini, R., and Pitzus, F.: Frequenza della talassemia in quattro popolazioni sarde e suoi rapporti con la distribuzione dei grieppi sanguigni e della malaria. (Frequency of thalassemia in four Sardinian populations and its relationships with the distribution of blood groups and malaria.) Boll. Ist. sieroterap. milanese 36: 206–218, May–June 1957.

(*10*) Smithies, O., and Hiller, O.: The genetic control of transferrins in humans. Biochem. J. 72: 121–126, January 1959.

(*11*) Blumberg, B. S., and Warren, L.: The effect of sialidase on transferrins and other serum proteins. Biochim. et biophys. acta. In press.

(*12*) Allison, A. C., and Blumberg, B. S.: An isoprecipitation reaction distinguishing human serum-protein types. Lancet 1:634–637, Mar. 25, 1961.

(*13*) Roberts, J. H. F.: Some associations between blood groups and disease. Brit. M. Bull. 15: 129–133, May 1959.

(*14*) Glynn, A. A., Glynn, L. E., and Holborow, E. J.: Secretion of blood group substances in rheumatic fever: A genetic requirement for susceptibility? Brit. M. J. 2: 266–270, Aug. 29, 1960.

(*15*) Clark, C. A., McConnell, R. B., and Sheppard, P. M.: ABO blood groups and secretor character in rheumatic carditis. Brit. M. J. 1: 21–23, Jan. 2, 1960.

(*16*) Kitchin, F. D., and Evans, W. H.: Genetic factors in thyroid disease. Brit. M. Bull. 16: 148–151, May 1960.

(*17*) Clements, F. W.: Goitrogens. Brit. M. Bull. 16: 131–137, May 1960.

(*18*) World Health Organization: Endemic goitre. Monogr. Series No. 44. Geneva, 1960.