ACTA GENETICAE MEDICAE ET GEMELLOLOGIAE

Volumen XV *N. 2 - Aprilis 1966*

World Health Organization

The Use of Twins in Epidemiological Studies

Report of the WHO Meeting of Investigators on Methodology of Twin Studies

Ricevuto il 21. II. 1966

WHO MEETING OF INVESTIGATORS ON METHODOLOGY OF TWIN STUDIES

Geneva, 26 October - 1 November 1965

Members:

Dr G. Allen, National Institute of Mental Health, Bethesda, Maryland, USA

Dr N. O. Borhani, Acting Chief, Bureau of Chronic Diseases, Department of Public Health, Berkeley, California, USA

Dr R. Cederlöf, Tvillingregistret, Institute of Hygiene, Karolinska Institutet, Stockholm, Sweden *(Rapporteur)*

Dr J. H. Edwards, Department of Social Medicine, University of Birmingham, England

Professor L. Gedda, Istituto di Genetica Medica, Università di Roma, Rome, Italy *(Vice-Chairman)*

Professor L. Friberg, Institute of Hygiene, Karolinska Institutet, Stockholm, Sweden

Dr B. Harvald, Bispebjerg Hospital, Copenhagen, Denmark

Dr M. S. Honeyman, Connecticut Twin Registry, Connecticut State Department of Health, Hartford, Connecticut, USA

Mr S. Jablon, Division of Medical Sciences, National Academy of Sciences - National Research Council, Washington, D.C., USA

Dr E. Jonsson, Institute of Hygiene, Karolinska Institutet, Stockholm, Sweden

Professor D. Klein, Institut de Génétique Médicale, Université de Genève, Geneva, Switzerland

Professor J. B. Lawson, Department of Obstetrics & Gynaecology, University of Ibadan, Ibadan, Nigeria

Dr H. Mitsuda, Osaka Medical College, Takatsuki, Osaka, Japan

Professor D. D. Reid, Department of Epidemiology & Medical Statistics, London School of Hygiene & Tropical Medicine, London, England *(Chairman)*

Mr J. Shields, Psychiatric Genetics Research Unit, Institute of Psychiatry, Maudsley Hospital, London, England

Secretariat:

Dr R. L. Kirk, Chief, Human Genetics, WHO

Mrs C. C. Standley, Scientist, Human Genetics, WHO *(Secretary)*

World Health Organization

The Use of Twins in Epidemiological Studies

Report of a WHO Meeting of Investigators *

The WHO Meeting of Investigators on Methodology of Twin Studies was held in Geneva from 24 October to 1 November 1965. The meeting was opened by Dr L. Verhoestraete, Director, Division of Health Protection and Promotion, who welcomed the participants. He said that this meeting of geneticists, epidemiologists and clinicians had been called to review the possible contribution of twin studies, a recognized tool in genetic research, to the broader field of epidemiological investigations, where so far they had found comparatively little use.

Professor D. D. Reid was elected Chairman; Professor L. Gedda, Vice-Chairman and Dr R. Cederlöf, Rapporteur.

1. Introduction

Studies of general population characteristics include vital statistics such as mortality and morbidity data which are obtained on a routine basis. These data constitute one of the best sources of raw material for epidemiological studies and provide easily accessible information for basic research and for programme planning in control of disease. Very often they raise questions concerning the epidemiology of chronic disease, which may help delineate the relevant factors and provide guiding hypotheses for further research in this field. A deeper understanding of disease may also be arrived at through the study of individual characteristics in cross-sectional, retrospective and prospective studies, and through experimental investigations. A cross-sectional study can be used not only in determining prevalence of disease in a population, but it can also serve as the base-point for a longitudinal study of chronic diseases.

* This report contains the collective views of an international group of experts and does not necessarily represent the decisions or the stated policy of the World Health Organization.

Unfortunately the chronic diseases are so complex in their nature that systematic inferences as to their etiology are difficult. Most statistical associations perhaps raise more important questions about many common diseases than they answer. Significant among these questions are " way of life ", social interactions, and a host of other factors which can be referred to as environmental.

In addition, it may be assumed that environmental factors will affect individuals differently according to their genetic make-up. It is apparent therefore that in the study of chronic diseases epidemiologists must go beyond their previous important, but obviously insufficient, approaches. They should begin to orient their thinking towards the contribution that population genetics can make to epidemiology. In fact, they must recognize that such epidemiological problems can be fully elucidated only if genetic factors are considered and studied concurrently with environmental factors.

Twins may offer particularly good opportunities for such studies. The two members of a pair of identical twins have the same genetic make-up and permit studies of response to different environments; on the other hand, the two members of a pair of non-identical twins are genetically different and can give some insight into the effect of similar environment on different constitutions.

Twin studies in the past have almost exclusively aimed at studying rather small series of twins for diseases and traits with a very strong genetic predisposition. They have been used only to a very limited extent in epidemiological studies. Now that larger twin series — covering whole populations — begin to be available in some countries the need for studying methods and uses for such data has become obvious.

2. Review of earlier and current uses of twin data

The first use made of twins in the study of hereditary phenomena dates back to the publication of Galton's work " The history of twins as a criterion of the relative powers of nature and nurture " in the year 1875. Between 1910 and 1940 the various European and American Schools contributed to the formulation and application of the theoretical " nature v. nurture " models.

During the last two decades twin data studies have been used extensively and found to be helpful in evaluating the role of genetic factors in the etiology of some diseases and in normal development. A large number of publications, particularly in the fields of psychiatry and malformations, have appeared on these subjects. A review up to 1950 is given by Gedda (1951) in his book " Studio dei Gemelli ".

However, up to now twin data have seldom been used in relation to epidemiological problems.

In order to get some information on present activities in the field of twin research with emphasis on studies of large series of twins, WHO carried out an *ad hoc* survey[1] on a world-wide basis by means of a mailed questionnaire to " twin researchers ". The inquiry concerned current or planned investigations and asked for details about, among others, the aim of the study, the type of data collected, the number of twins, the methods of sampling, zygosity determination and published results. The institutions to which the questionnaire was mailed were chosen after consultation with different units within WHO and with experts outside the Organization. The information obtained cannot, therefore, be regarded as fully representative. Nevertheless, the results appear to present a useful picture of twin research today.

Tab. 1 contains some essential data condensed from the survey. Extensive twin studies are going on in several European countries, in the USA, as well as in Japan.

The table shows that out of 39 reported twin series, covering in all about 60 different ongoing or planned research projects, about two thirds are of the " selected " type, i.e. obtained from hospitals, clinics, schools, etc. The remaining registries cover whole populations or well-defined parts of a population; about one third of these, however, contain data obtained from birth records only. The twin series vary in size, from 40-50 pairs to more than 20,000; the smaller series mostly consist of case-studies and the larger include complete populations.

The subjects covered also display a wide range, although the majority of programmes concern psychiatry, psychology and normal development. Chronic disease — in a few cases with special emphasis on effects of different environmental agents (e.g. smoking, air pollution) — has been the subject of about 10 investigations, half of which are based on series of more than 2000 twin-pairs. Investigations on twinning rate, basically of a statistical nature, have been performed in about 10 series. Other subjects have been the recording of physiological data and studies on malformations as well as general mortality.

Twin series of the non-selected type and large enough for use in epidemiological investigations are available mainly in the USA and in the Scandinavian countries. Some of these include current addresses of twins, others do not at present but comprise a large number of twins and must be looked upon as very important potential sources for epidemiological research.

[1] Copies of the "WHO International Survey of Twin Registers and Studies" are available free of charge on request to the Human Genetics Unit, World Health Organization, Geneva, Switzerland.

Tab. 1

Subject of study	Type of twin material *	Number of twin pairs			
		-500	501-2000	2001-10000	10001+
Chronic disease	Selected twins	2	2		
	Non-selected twins A		1	2	1
	Non-selected twins B				3
Psychiatry, psychology, normal development	Selected twins	6	9	3	
	Non-selected twins A		2	2	1
	Non-selected twins B				1
Twinning rate and similar problems	Selected twins	2	1	1	
	Non-selected twins A		1		1
	Non-selected twins B				4
Recording of physiological data	Selected twins	1	1	1	
	Non-selected twins A			2	
	Non-selected twins B				
Malformations	Selected twins			1	
	Non-selected twins A			1	
	Non-selected twins B				
Mortality	Selected twins	1			
	Non-selected twins A				1
	Non-selected twins B				2
Non-specified	Selected twins		1	1	
	Non-selected twins A				1
	Non-selected twins B		1		
Other	Selected twins	3			
	Non-selected twins A	2			
	Non-selected twins B				

* Selected twins, mainly from hospitals, clinics, military registers, etc.
Non-selected twins A: current address known
Non-selected twins B: current address unknown, mainly birth data
N. B. The same investigation may be recorded under more than one "subject of study" heading.

3. Methodological problems

1. General Aspects of the Twin Method

This section will present some general considerations on the evaluation and interpretation of data from twin-studies. They concern the particular nature of twin data from the physiological point of view and the question of the extent to which environmental differences between identical twin partners (MZ) [2] really are of the same magnitude as for non-identical twins (DZ).

[2] The group recommends that the term "monozygotic" be used in preference to monozygous or monovular, which have in the past led to confusion. Similarly, "dizygotic" instead of dizygous or binovular.

Twins are peculiar in that they develop in a uterus primarily adapted to the nourishment of a single foetus. They are usually less mature at birth when they are exposed to an increased risk of obstetric complications. If both survive, their development may be impaired by difficulties in lactation. Their upbringing is unusual as it is shared by a sib of their own age and they are known to have an increased mortality in childhood. Some differences in disease experience between twins and singletons may merely reflect an increase in the mean age and parity of mothers of dizygotic twins, or, in groups including different races, differing liabilities to both twinning and disease.

The use of twin data to make inferences about normal variation or disease susceptibility in singletons may therefore be misleading. However, it would seem reasonable to regard the peculiarities imposed by gestation and birth as less influential as life advances, so that, although serious misinterpretations might arise from studies of disorders of foetal life and infancy, and to a smaller extent, of childhood, no serious bias need be anticipated in relation to most disorders of adult life.

Large twin studies are most commonly resorted to in epidemiology in the hope of discovering the causes of poorly understood diseases, i.e. diseases showing no simple Mendelian pattern and no rigid association with an obvious environmental agent. All diseases of unknown cause can be presumed, without investigation, to have at least some interdependence of genetic and environmental factors; exclusive dependence on genetic mechanisms will not be demonstrable except in the event that virtually complete concordance is observed in MZ twins, accompanied by a significantly lower concordance rate in DZ twins. Hence, with respect to genetic determination, a comparison of concordance in MZ and DZ twins can do no more than to draw attention to the presumptive importance of genetic factors.

One may even carry the argument one stage further: a study of twins raised in the same household cannot, in principle, exclude complete environmental determination of the disease under investigation, because normal MZ twins can be presumed, in general, to share more features of environment and experience than do DZ twins. As long as the etiology of the disease is not fully defined, the possibility cannot be excluded that some item of experience shared by MZ partners is a determining factor. As a theoretical argument, this is almost impossible to refute. In reality, most shared post-natal experiences of MZ twins are probably not qualitatively different from those shared by DZ partners or even by sibs. Consequently a critical environmental factor that could explain a relatively high concordance in MZ twins would probably raise the concordance in DZ twins well above that of other

sib pairs, since DZ twins share a common environment to a greater extent than sibs do. Therefore the finding of a concordance in MZ twins that is higher than the population incidence, coupled with a lower concordance in DZ twins, is suggestive of a genetic influence.

The evidence would be strengthened if, as may be the case, twins raised in different homes tend to be as alike as those raised in the same home. However, separated twins are at present too infrequent and atypical to be relied upon as a major source of evidence for genetic effects. With respect to a disease showing a raised concordance in sibs, even a study of sibs separated in infancy can serve to differentiate genetic from environmental factors.

For the elucidation of possible *environmental* causes of diseases, the shortcomings of twin studies are not theoretical but practical. Discovery of critical environmental differences between a series of affected MZ twins and the matching series of their unaffected partners is theoretically possible, but often proves quite difficult. Another approach has great promise, but has up to now been used on a very limited scale: one may first select a large series of MZ twins who differ in their exposure to some suspected environmental circumstance, and then follow them to see if subsequent morbidity is related to exposure. This use of twins can be considered as analogous to an experimental design. If the differences between MZ pairs in the study were small, the suspected agent would be unlikely to be of major importance. The limitation to this approach is that in certain circumstances MZ twins may differ from the normal population in such a way that their disease experience is peculiar.

The experimental co-twin control method provides yet another approach to study of environmental agents of disease. While one would not, of course, risk exposing randomly selected partners in a series of MZ twins to a serious disease, one could design a prospective inquiry in which one member of each pair would be provided with a diet or other treatment thought to protect from the responsible agent.

2. Zygosity Determination

There appear to be four distinct mechanisms of twinning in man, three of which lead, barring accidents in the apportioning of the chromosomes during early cell divisions, to genetically identical twins. Although all monochorionic and the rare monamniotic twins are MZ, it appears from Tab. 2 that in a Caucasian population up to one third of MZ twins, those which result from doubling of the blastocyst, may be dichorionic.

This table refers to a consecutive series (Cameron *et al.*, 1965) of 581 twins in which placentation and sex were known, and in which extensive tests on zygosity were carried out in 490. 90% of twins with 2 choria were estimated from blood and enzyme grouping to be DZ; 100% of those with 1 chorion were MZ. It should be mentioned in this context that if at birth a piece of the membrane between the

Tab. 2. The four mechanisms of twinning in man

	DZ	MZ		
	Double ovulation	Doubling of blastocyst	Early doubling of embryonic disc	Late doubling of embryonic disc
Proportion in Birmingham series	72%	8%	17%	3%
Number of amnia	2	2	2	1
Number of choria	2	2	1	1
Number of zygotes	2	1	1	1

foetal sacs is sectioned, the evidence for zygosity can be preserved indefinitely as a histological preparation, a point of practical importance in relation to " spare part " surgery.

Diagnosis of zygosity based on counting the placentae is much less reliable. In the series referred to above, when placentation was used as a criterion, 90% of twins with 2 placentae were found on serological diagnosis to be DZ but only 40% of those with one placenta were MZ.

Zygosity can be determined by considering similarities between twins. This can be done with great accuracy by multiple tests of blood groups, serum protein variants, and red-cell and placental enzyme systems, which in a Caucasian population can distinguish zygosity in over 95% of cases. Any pair differing in only a single system should have the test repeated. When possible, testing should be done on both twins at the same time. Occasional errors can lead to misclassification, mainly through monozygotic twins being classified as dizygotic as the result of a single erroneous observation, a bias which increases with the number of systems studied.

Zygosity may also be ascertained by comparing more complex traits such as facies, hair colour, finger and palm prints, size, voice-sounds, etc. Some of these have proved to be of a high validity provided the characteristics are not disturbed by the condition under consideration. Details of calculating the chances of twins being MZ, after a given number of tests, may be found in Maynard-Smith *et al.* (1961).

Diagnosis of zygosity by means of questionnaires has proved to be very reliable. In studies in Sweden and the USA, the question: " When growing up were you as alike as two peas in a pod or of a family likeness only? " could distinguish MZ from DZ adult twins. Where both members of the twin pairs had answered this question consistently, over 95% were correctly identified on the evidence of blood and serum grouping. In Tab. 3, some detailed results are given from the Swedish Investigation. This question is also used in studies in Italy.

Tab. 3. Reliability of zygosity diagnosis according to questionnaire method (Cederlöf *et al.*, 1961)

	Blood diagnosis		Percentage of total no. of MZ/DZ according to questionnaire method	Percentage of total no. of MZ/DZ diagnosed by blood examinations
	MZ	DZ		
MZ: Both twins replying " as alike as two peas "	71	1	99	81
DZ: Both twins replying " only family likeness "	10	104	91	93

The questionnaire method, however, should not be used on small numbers, for studies on babies or young children, or when other methods are applicable and feasible. They are usually inappropriate when it is important for every single case to be correctly diagnosed.

When zygosity cannot be studied in detail in large population groups, the Weinberg " differential method " yields, on the evidence of sex alone, a useful estimate of the proportions of twin types (Weinberg, 1901).

There are advantages in using several methods for diagnosing zygosity, as these are cumulative and largely independent, and can be applied serially until an adequate decision is reached. The optimal order of testing is dependent upon the circumstances. Accurate blood analysis is a method of choice, provided due attention is paid to the consequences of errors. It may be mentioned that an additional advantage of determining blood groups and other markers is the information that may be obtained on genetic linkage and on the associations of blood group and other markers with disease or other types of variation. For studies of large groups, however, such methods are often impracticable, and simpler procedures will usually be adequate. There is a great need for further research in the comparison of various methods in terms of cost, acceptability and precision.

3. Types of Twin Samples

There are at least five types of twin samples. Since the method of sampling determines the nature and reliability of any conclusions that may be drawn, the different types are discussed below, although they are not all suitable for epidemiological research:

1. *Single case reports.* A single pair of twins, very carefully studied, may be useful in showing what *can* occur, and may therefore justify publication. Detailed study of a single pair of twins, when one twin presents the disease, may assist in the diagnosis of a less affected member or may allow the prevention of disease by a modification of the environment.

2. *Multiple case reports.* The reporting of several twin pairs having the same condition is more informative than single case reports. However, while the analysis of multiple case reports may give a spurious impression of valid sampling, the biases of individual case reports remain, as the source population cannot be defined for purposes of statistical control or correction.

In regard to case reports, one may expect higher concordance than in series based on total ascertainment of a twin population as there may be a tendency to report concordant rather than discordant pairs. Bias of this sort might be reduced if physicians' reports of rare conditions in twins were accumulated in a central registry.

3. *Volunteer series.* Some studies have depended upon twins who responded to public appeals. This category includes any sample supposedly drawn from the whole twin population, but for which representativeness cannot be assessed.

4. *Consecutive series.* Many valuable twin studies have obtained twins from a defined population of clinical or hospital patients. This method does not eliminate all the biases of single case reports, but it does provide a reference population and permit assessment of bias.

5. *Unselected samples of twins.* The ideal twin series would include all the pairs in a population. Since it is not practical to obtain information on such a series, it is usually necessary to use samples. There are two ways in which selection may be accomplished: by drawing a quasi-random sample of the whole population, such as the population of births selected by time or place, or persons entering military service, etc. Some of these methods of choosing the sample to be studied actually imply selection of a kind that may limit the applicability of results and the investigator must be sure that the limitations do not affect the question that he is attempting to study. Unselected samples are economic only in the study of common dis-

ease, but do have the great advantage of excluding most of the causes of bias considered above.

In many twin studies it is required that both members of each pair should be observed. This will result in a great number of living individuals being left out of the sample simply because their partners are dead. Since there is an intra-pair correlation with respect to mortality, this selection would exclude physically weaker individuals from the sample and make it healthier. Secondly, in twin studies concerned with environmental factors and health, the selection of twin pairs may be made on the basis of their being discordant for exposure. As there very often exists a strong concordance also with regard to environmental factors in twin pairs, those who are selected as being discordant might be unusual individuals.

4. Organization and Maintenance of Twin Registers

The aim of registration is to obtain a sample of twins that is both representative and large enough to yield reliable data, ideally the whole twin population of a state or country.

Routine systems of registration of vital events and of population enumeration exist in almost all nations. Although these systems are created primarily to serve administrative purposes, they have very valuable research uses. Twin rosters or lists can be derived from them, and they can serve to trace members of the roster and obtain information on events of interest such as disease or death.

In many countries records of births and still-births specify whether the birth is a multiple one. Registers based on the routine search of such birth records have the great advantage of being complete and hence unbiased. Twins may also be identified from lists of military recruits. These have the defect of applying wholly or chiefly to males, and including only persons who survive in good health to the age of military service.

Methods of constructing twin registers, other than on the basis of complete enumerations, may produce rosters that are biased. One such method is to register the twins on the basis of hospital records, e.g. at maternity wards if the procedures are such as to guarantee that all twin births are noted. A rather different procedure is to seek twins among in- and out-patients seeking medical care at medical and psychiatric clinics. However, even assuming that complete ascertainment of twins can be obtained in this way, conclusions cannot necessarily be generalized to the whole population unless those seeking care are typical of all affected persons.

Another method of compiling a twin register is the so-called " mass media

method " requesting all twins to register by radio, television or newspaper notices. It is also possible to distribute simple questionnaires to all households within an area and ask the twins living there to return the questionnaire. Such methods, which depend upon volunteers, may be supplemented with incentives of various kinds, including the possibility of receiving care and help from specialized institutes.

There are two aspects to the maintenance of a twin register. One is to keep track of the twins' adresses, and the other is to obtain current information in regard to mortality, morbidity, socio-economic status or other desired information. Whenever studies of the environment are carried out on twins, it is recommended that, because of the expense and difficulty of collecting data, information on as many of the environmental factors as possible be collected and studied at the same time. While the specific characteristics of record systems that can be useful vary in different countries, some possibilities of common applicability can be mentioned.

1. *Death records.* While death certificates do not ordinarily contain the information as to whether the deceased was a member of a twin pair, where it is possible to match a roster against the list of recorded deaths, this is the method of choice for the study of mortality among large groups.

2. *Censuses.* It is not usual for census enumerations to include the fact that the person enumerated is a twin. However, where the possibility of matching exists, use of census lists provides an expedient method of obtaining recent addresses for the members of a large roster.

3. *Other administrative files and registries.* Administrative practices vary widely among the countries. A variety of resources may be available as aids in tracing and data collection:

a) lists of military recruits
b) population registries
c) family registers
d) registers of persons admitted to hospital
e) insurance and social security records
f) postal facilities
g) tax records.

It should be emphasized that the ease or even possibility of using these resources is dependent upon record linking from one administrative file to another. Only if the linking problem is solved can these files be of value for medical and epidemiological research. This applies to medical research in general, not only to twin studies. The solution of this problem is well within the capacity of current computer

techniques and it is extremely important to encourage the relatively small legislative and administrative changes required to facilitate this. It is highly desirable to update files as much as possible by record linking to routinely maintained files, since information so derived has the advantages of completeness and low cost.

For some purposes, however, it may be necessary to establish direct contact with the twins. For large series, such contact would be economically feasible only by mail. Questionnaires, when used, must of course be well-designed and standardized. There are a variety of standardized questionnaires which, experience has shown, can be adapted to use by mail, for example the questionnaire regarding angina pectoris, which was developed at the London School of Hygiene and adopted by the WHO for use in field surveys, (Rose, 1962 and Pearl *et al.*, 1965) and the British Medical Research Council questionnaire regarding chronic bronchitis.

One problem that must be considered in relation to mail questionnaires is the possibility that some pairs of twins will fill them out in consultation with one another. To the extent that this occurs, the validity of the survey will be decreased. Attention must also be paid to other aspects of the validity and completeness of information obtained by mail.

5. Presentation and Analysis of Twin Data

One method in common use in analyzing and interpreting categorical twin data is the calculation of concordance rates among monozygotic and dizygotic twins and, based upon these rates, a heritability index,[3] indicating the relative influence of genetic and environmental factors. These concepts might be useful also in interpreting data from epidemiological twin investigations, but should be complemented by other more straightforward methods permitting comparison between twin series where the prevalence of disease might be different.

In the following, the concept of concordance will be first commented upon, some different ways of calculating the rates will be presented, and, lastly, some examples given of other useful methods.

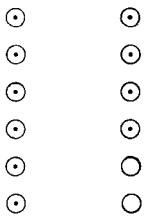
Certain complications arise in applying classical twin methodology to the study of disease, especially those diseases which occur late in life. The two members of a pair will hardly ever acquire the same disease at exactly the same time, so that most

[3] There was general agreement among the group that the heritability index, combining as it did the concordance rate for a specified trait among MZ twins with the concordance rate for the same trait among DZ twins, was difficult to interpret.

pairs in which the disease occurs will be initially discordant. The second member of the pair may die of a cause unrelated to the disease, e.g. an accident. It is apparent, therefore that the use of concepts such as " the proportion of concordant pairs " which are appropriate to the study of such conditions as colour-blindness, are not wholly adequate for the study of coronary disease. One way of approaching this problem is to study age and sex-specific morbidity or mortality rates within appropriate subgroups, if the sample is large enough to support such detailed analysis.

Pairs of twins including an affected individual may be collected either by ascertaining twins and rejecting those who are both unaffected or by ascertaining affected individuals and including only those who are one of a twin pair. The proportions of pairs with two affected members (the concordance rate) will differ in these two procedures since, in the latter, some pairs may be ascertained twice and will, therefore, be counted twice. As an example, consider a population containing 40 pairs of which ten are concordant. If ascertained through twins the concordance rate might be said to be 10/40 or 25%. If ascertainment were by affected individuals, and complete, 50 pairs, 10 of them in duplicate, would have been ascertained and the concordance would be 20/50 or 40%.

This may be illustrated diagrammatically by reference to a model twin population of six pairs (for simplicity, only MZ), in which four are concordant with respect to a disease and two are discordant:



For this population the concordance rate may be stated in two quite different ways. In terms of pairs, the rate is 4/6 or 66%. In terms of affected persons having affected partners, the concordance rate is 8/10 or 80%.[4]

If data are to be analyzed in terms of concordance, the criteria must be clearly stated including, where relevant, age of onset and length of follow-up of unaffected partners.

[4] The latter method has the advantage that the concordance rate is the same as the probability of the co-twin being affected.

Another way of evaluating genetic influence on the basis of concordance could be referred to as a " coincidence rate ". It is useful when handling data from different twin series with different incidence rates for a certain symptom. This may be the case when comparing studies from different population groups with different exposures, whether to a single factor, e.g. tobacco smoke, or to the entire environment, e.g. different geographical areas. The method requires that information be obtained from all pairs in the target population or in a random sample thereof, including twin pairs with neither partner affected. As with other measures of concordance, comparison between groups must be made with reference to the population frequencies of disease. Age and sex-specific morbidity rates for the whole twin population studied are applied in turn to each contrasted group, for example monozygotic or dizygotic sets of twins, to obtain the expected frequency of the coincidence of disease in such sets on the basis of genetic independence. Comparison of the observed and expected rates will then show whether the excess in observed concordance over the frequency expected, for example in monozygotic sets, is greater than the corresponding excess in dizygotic pairs.

The analysis of graded or metric traits may be undertaken in terms of either intrapair correlations or variance analysis, and standardization of the observations is sometimes appropriate, e. g. with regard to sex, age, smoking habits, etc. Whatever method is chosen, sufficient data should be provided to permit either kind of calculation, preferably on the raw data, as well as on the standardized ones.

6. Other Methodological Problems

The possibility that twins may in certain respects differ from singletons has been mentioned repeatedly in the foregoing sections of the report. Knowledge of the extent to which this may be true, however, is very limited and should be increased. When dealing with morbidity and mortality rates in twins, it would be most valuable if the corresponding figures in the normal population — of course with reference to sex and age — could be given. The twin sample — whether randomly chosen or not — should also be described in regard to as many relevant background variables as possible. This would help the investigator to judge — at least approximately — in what respects and to what extent the sample might differ from other twin samples reported or from the population as a whole.

Twin studies are a particular example of paired-relative studies, and there will sometimes be advantages in extending the study by the inclusion of other relatives,

in particular of sibs. The sibs thus constitute a valuable control material of singletons to be compared with the twins. Sibs are of particular relevance since differences between dizygotic twins and other sib pairs are the only measures available of the influence of the common pregnancy and unusual upbringing peculiar to twins.

The effect of an environmental factor on some aspect of health can only be studied in a population group in which there is an association between the factor and the suspected effect. For reasons already mentioned, twins may differ in health experience from the population in general and may not exhibit the associations between specific factors and disease which are apparent in the latter. One way to ensure that the association between a specific environmental factor and disease, which one wishes to study, is present in unrelated members of these twin populations is randomly to choose *one* twin from each pair, to form a " new " population of individuals with no more in common, from the genetic point of view, than a normal population, and seek evidence for the association in the " new " population. The reason for using one twin only of each pair is statistical. Most significance tests require that observations be unrelated; this is not the case if both twins of a pair are included.

4. Use of twin data for study of genetic and environmental effects

It is the opinion of the Group that studies of twins can provide useful information on the relative contribution of genetic and environmental components in disease. Such information is needed especially for the chronic diseases such as malignancy and cardiovascular disease. The results of epidemiological study of these conditions appear to indicate that they are of complex etiology, and it seems likely that the influence of environmental factors in their causation varies according to the genetic material on which they act.

Only a few twin studies concerning chronic disease of late manifestation have so far been published, probably because of the inherent difficulties of gaining materials of sufficient numbers and quality. The twin registries which, in recent years, have been or are being established in different European countries, as well as in the USA, open the possibility of overcoming these difficulties. Two main kinds of twin studies should be distinguished:

1. Studies aiming at comparison of the occurrence of diseases in representative pairs of monozygotic and dizygotic twins.

2. Twin control studies, wherein the two members of a genetically identical pair are exposed to differing environmental influences.

Studies of the first type are based on the assumption that since monozygotic twins are genetically identical, all within-pair variations in such twins are due to environment; whereas in dizygotic twins both genetic and environmental factors may be responsible. Furthermore, it is assumed that the within-pair environmental differences for like-sexed dizygotic twins are not substantially greater than for MZ twins. This condition is hardly ever totally fulfilled and in all comparisons due regard must be given to the possibility that observed differences might be caused by factors of this kind. Studies of other sib-pairs, and if possible MZ pairs reared apart, might shed some light on this problem.

The second type of twin studies can be considered as analogous to an experimental study, aimed at disclosing a possible causal relationship between environment and disease. For example, the influence of cigarette smoking on mortality and morbidity can be studied by following a group of MZ twins in whom one member is a smoker and the other is not.

The presence of interaction between genetic and environmental factors may sometimes be brought to light by a study of MZ twin pairs in which both or neither partner has been exposed to some environmental factor such as smoking. Estimates of the genetic contribution to response may then be derived from a comparison of the number of doubly affected pairs found and expected on the basis of the age, sex and other specific rates prevailing in the twin population studied.

There are other examples of environmental factors that might suitably be studied by means of twin investigations. For example the different morbidity in urban compared with rural dwellers might well to some extent be of a genetic origin. Further, studies on twins on the effect of air pollution would provide the investigator with an ideal control material.

It is the opinion of the Group that there exist a number of environmental conditions that are of great importance from the public health point of view and well suited for study on twins, for instance occupational diseases. For a number of them, e.g. silicosis, byssinosis, chronic bronchitis, it is not known to what extent hereditary factors play a role. It is highly desirable that studies on twins in these fields be carried out.

Comparative analysis of environmental differences in discordant MZ twins could reasonably be proposed for many pathological conditions such as coronary heart disease, hypertension, peptic ulcer, diabetes, among others, but it should be made

clear that this sort of study necessitates samples of considerable size; until now no such study has been published, probably because a sufficient number of twins could not be obtained.

In addition to the types of twin-studies mentioned above, the Group recognizes the usefulness of studies of twinning phenomena, secular trends in twinning and comparison of twins with the general population in regard to mortality, morbidity, fertility, etc. Studies of MZ twins, especially those reared apart, will yield information on environmental effects on the evolution as well as the causation of disease. Furthermore, in special cases of disease with a strong genetic component where a high degree of concordance is found in MZ twins, the study of hitherto unaffected co-twins permits observation of the presymptomatic development of the disease, e.g. in diabetes mellitus, or in schizophrenia. It should in this context be stressed that detailed clinical study of each member of the twin pair is of particular importance in psychiatric studies.

Twin studies in schizophrenia and other psychiatric disorders, which have been useful in the past, will no doubt continue to be made. There is value here in diversity of approach rather than in pooling of data. Studies of MZ and DZ twins together with other family members can with advantage be undertaken in a variety of settings including long-stay and short-stay hospitals. Parallel studies in countries where they have not previously been carried out would be desirable.

The reporting of representative series of discordant and partially discordant MZ pairs is of particular importance in psychiatry. The investigation of such pairs by clinical, psychometric and other appropriate means, and their follow-up, may shed light on the variability of manifestation in the co-twin, the environmental reasons for the difference and the possibility of etiological heterogeneity.

5. Need and possibilities for international collaboration

The incidence of arteriosclerotic heart disease and cerebrovascular disease differs widely between countries. Efforts to determine the genetic and environmental factors involved in these as well as other chronic diseases would be enhanced by studies of twins in parts of the world where contrasts in findings due to different gene-pools, or different environments, may be informative.

Collaborative twin studies will also be valuable when data from any one study is insufficient to give definitive results in a reasonable time, which may be the case even for common diseases, especially when the sample is broken down in diagnostic

subgroups or by age and sex. Breakdown into diagnostic sub-groups would seem to be a prerequisite to progress in the study of cardiovascular diseases or cancer. Parallel studies with the same aims and with as similar design as possible will, if indicating similar trends, considerably strengthen confidence in the results.

When studying relatively rare traits or traits occurring only in small groups of a population it might be possible to pool data from different investigations. As twins reared apart or migrant twins are very rare in any one twin-registry, data from such twins would also be easier to assess if they were pooled on an international basis.

What has been said above concerning the need for collaboration also emphasizes the necessity of a certain standardization of methods.

To make the above-mentioned studies possible the Group recognizes the need for some centralized agency, such as WHO, to promote and co-ordinate international collaboration in twin studies. Such an agency could act as a clearing-house of information for such endeavours, convene meetings of interested and suitable workers, assist in the design of studies and provide facilities for training junior work ers and exchange of personnel. It may also be desirable for data from these studies to be collected and analyzed centrally.

Bibliography

Cameron A. H., Edwards J. H., Wingham J. (1965). *Unpublished.*

Cederlöf R., Friberg L., Jonsson E., Kaij L. (1961). Studies on similarity diagnosis in twins with the aid of mailed questionnaires. *Acta Genet.* (Basel), **11:** 338-362.

Gedda L. (1951). Studio dei Gemelli. Ed. Orizzonte Medico, Roma.

Maynard-Smith S., Penrose L. S., Smith C. A. B. (1961). Mathematical tables for research work in human genetics. Churchill, London.

Medical Research Council's Committee on the Aetiology of Chronic Bronchitis (1960). *Brit. Med.* J., ii, 1665.

Pearl R. B., Levine D. B., Gerson E. J. (1965). Studies of disease among migrants and native populations in Gt. Britain, Norway and the United States. II. Conduct of field work in the United States. *Nat. Cancer Inst. Monogr.*, **19:** *In press.*

Rose G. A. (1962). The diagnosis of ischaemic heart pain and intermittent claudication in field surveys. *Bull. Wld. Hlth. Org.*, **27:** 645-658.

Weinberg W. (1901). *Arch. Ges. Physiol.*, **88:** 346-430.