Some Problems in Population Genetics

CARL HENRY ALSTRÖM*

It has been known for a long time that patients suffering from rare diseases following a recessive mode of inheritance are, in a high proportion of cases, offspring of marriages between related partners who themselves are healthy. In such cases varying percentages of 'consanguineous marriages' have been reported. Such figures are useful as a general orientation. The great differences found between different reports reflect, however, the extent to which genealogical research into the pedigrees has been carried out, and it is the rate of first-cousin marriages among parents of systematically collected probands which is of the most practical importance. Only where data from first-cousin marriages can systematically be collected for the subject under investigation and compared with the rate of such marriages in the general population is the information of value.

Garrod (1901, 1902) seems to have been the first clinician to observe that a high percentage of parents of patients with rare recessive diseases were first-cousins, since such marriages increase the probability for heterozygotes to come together, first-cousins having one-eighth of their genes in common. Furthermore, Garrod's clarifying discussions with biologists, for instance with Bateson, in these early days of rediscovery of the Mendelian laws do not lessen the greatness of Garrod's pioneer work; in fact the contrary is the case. It is Garrod's sharp observations combined with his theoretical reasoning that solved the puzzle and his concept of inborn errors of metabolism in this connexion was far ahead of his time. After Garrod many British clinicians have carried the fine tradition in human genetics further, and in ophthalmology, which has contributed so much to human genetics, in our day Sorsby has an international leading position as a clinical geneticist.

The German physician Weinberg (1920, p. 40) seems to have been the first to deduce a correct formula for the relation between the prevalence of

first-cousin marriages between the parents of recessive affected individuals (k), the prevalence of a recessive gene (r), and the prevalence of first-cousin marriages in the general population (c):

$$k = \frac{c(1+15r)}{c(1-r)+16r};$$

a formula applicable on the assumption of panmixia, i.e. mating at random. Later on Dahlberg (1929) made the same deduction of the relation between the three variables.

It is difficult to say to what degree 'mating at random', with regard to first-cousin marriages, occurs in human populations. Social factors of different kinds certainly interfere. The prevalence of such marriages seems to vary in European countries, usually between 0.5 and 2% roughly. The minimum estimate for England of 0.61% made by Julia Bell (1940) is reasonable. The critical argument by Dahlberg (1943, p. 89) that the value is too high because it was obtained from patients treated in hospitals does not seem to be valid. The estimate was made on large general hospital populations, and one must agree with Bell that it is safe to conclude that her figure is an underestimate of the rate for the country as a whole. She gives several reasons for this, one being that hospitals with chronic patients where we can expect to find hereditary cases are not included in the study. Of interest is the fact that the rate for London hospital population was 0.65% first-cousin marriages. In Sweden some investigations from different parts of the country have been published, indicating direct estimates above 1%. On the basis of investigations of three different, rare recessive diseases, Sjögren (1931), Alström and Olson (1957), and Hallgren (1959), using Weinberg's formula, came independently to an indirect estimate of about 1.5% first-cousin marriages as an average for the whole population of Sweden. All three diseases are scattered over the whole country, with somewhat different prevalences in different areas, and it is of interest to notice the good agreement when using Weinberg's formula.

For Switzerland, Ammann, Klein, and Böhlinger

* Address: Karolinska Institutets Psykiatriska Klinik vid S:t Görans sjukhus, Box 12500, S-112 81 Stockholm, Sweden.

(1961) have given a rate of 2% for consanguineous marriages (presumably first-cousin marriages). In an area including 5 cantons they found the prevalence of 'dégénérescence pigmentaire de la rétine du type récessif autosomique' to be 14.8 in 100,000. The authors are using an approximation of Weinberg's formula:

$$k = \frac{c(1 + 15r)}{16r};$$

given by Dahlberg (1948, p. 63). They insert their estimated value of the gene frequency, the assumed value for consanguineous marriages in the general population, and with the help of the formula arrive at a calculated 'fréquence des mariages consanguins' of 12.2% among the parents of the probands. This was about half what their investigation had shown, namely 26.7%, and for that reason they conclude that, 'Cet écart montre que la formule de Dahlberg ne peut être utilisée qu'à titre d'indication générale et ne saurait remplacer une enquête personnelle.'

This conclusion and calculation are based on a misinterpretation of the formula used, which is relevant in regard to first-cousin marriages and under the assumption of panmixia (mating at random in a numerically stable population). In fact Ammann *et al.* (1961) in their investigation found a rate of 9.9% first-cousin marriages (see p. 113 in their publication), so one must say that the agreement between observed and calculated values is good. If Weinberg's formula (see above) is used, the calculated rate will be 11.1% and the agreement is still better. If one assumes a rate of 1% for marriages between first-cousins in the general population in Switzerland, Weinberg's formula gives the value of 5.9%, and even that is not in bad agreement with the observed value.

If, however, the factual rate of first-cousin marriages in the general population in Switzerland is about 1%, the authors observed an increased rate of such marriages against the calculated rate among the parents of their probands—9.9% against 5.9%—which may indicate that their material is heterogeneous. The high prevalence of the recessive affected which they report (not less than 14.8 in 100,000) may also indicate that their material could be heterogeneous.

Broadly speaking it seems, however, as if Weinberg's formula is useful. One conclusion from the population-genetical viewpoint is that panmixia may be presumed to exist in many, perhaps most, European countries. In some countries, however, the first-cousin marriage rate is extremely low in the general population. This seems to be the case in

Finland, where in two interesting investigations of rare, recessive diseases, Norio (1966) and later Nuutila (1968) found none or very few first-cousin marriages between parents of the affected. The negative attitude in Finland against marriage between first-cousins seems to be a conservative tradition since the time of union with Sweden, when the legal and above all the clerical authorities were very much against it. This was in contrast to Denmark and most European countries at that time. Only in 1829 was the law in Sweden relaxed and the resistance from the church broken down; since then there seems to have been a steady increase in the rate of marriages between first-cousins (Alström, 1958).

Under the assumption of panmixia, i.e. mating at random in a numerically stable population, Dahlberg (1929) has with the aid of the rate of first-cousin marriages given a formula for estimating the size of the 'isolates', i.e. the smaller units of subpopulations of which a large population, a nation, consists:

$$n = \frac{2b(b-1)}{c};$$

where (b) is the average number of children who grow up and marry in each sibship, under the assumption of stable population 2 in number, (c) is the rate of first-cousin marriages in the general population, and (n) is the size of the isolates (see for instance Dahlberg, 1948, p. 91). This approach to the problem was a pioneer work at that time and a valuable contribution to the problem. A rate of 1.5% first-cousin marriages in the population will correspond to an isolate size of about 270 individuals. This gives a considerable random local genetic differentiation if the isolate is of the 'island' model. The latter is not the case in European countries where the subpopulations are continually distributed, even with large variations in density of the population. The first-cousin marriage rates are not useful in this connexion. When the main isolating factor is that of distance, a method devised by Wright (1951) for estimating the 'isolate' or 'neighbourhood' size has proved useful even for human populations. When applied to Swedish subpopulations for the years 1800–1824 it is shown that the 'effective population size' of the 'neighbourhoods' usually exceeds 1,000 individuals (Alström, 1958, p. 328). No true genetic differentiation can consequently be anticipated for about 90% of the Swedish population, even at that time. For the 10% of the population living in the most northern part of Sweden that may, however, have been the case (Alström, 1963).

Another factor, studied in a subpopulation in

FIG. Basic families distributed according to number of children married and having offspring.

the central southern part of the country, may in this connexion be of interest. For the population in a small parish M-a the 'effective population size of the neighbourhood' could be calculated to be about 1,100 individuals during the period 1800-1824. The inhabitants in the parish were farmers living in a wooded and agricultural area. During the years 1800-1824, 365 marriages, representing 730 partners, were contracted. In the Figure on the upper line are the proband families, generation I of the study, distributed according to the number of children (generation II).

In the marriages of generation I, 1,112 children were born altogether. Of these, 615 attained adult age and married and, in their turn had children, i.e. transmitted the genetic information obtained from the parent generation further. These 615 represent the fertile part of generation II and are in the figure distributed along the lower line according to their origin from families of generation I. All the married individuals of generation I are followed from the time of contracting marriages to death, and all the individuals of generation II are followed from birth to death.

It is shown in the Figure that 21% of the proband families, i.e. generation I, have not contributed to the next generation, i.e. they have disappeared from the gene pool of the population. It is also evident from the Figure that one half of generation I had two or more children with fertile marriages, and that this half of the first generation is responsible for more than 80% of the next fertile generation. Furthermore, a quarter of the first generation, those with three children or more, is responsible for more than a half of the next fertile generation, i.e. that part of the gene pool in the population that is transmitted further. (A reservation, however, must be made for the illegitimate children who have not been included in the study.)

Contracting marriage after fertile age, or a short duration of marriage due to the death of one

partner, may explain these facts. The families belonging to generation I, distributed along the upper line in the Figure, have been tabulated below according to the proportion of female partners married before 35 years of age (i.e. married during the fertile period of life) and according to the duration of the marriage; exceeding 2, 4, 6, and 10 years.

At that time no effective birth control existed, especially in this pronounced rural area. The family groups in generation I with 2, 3, and 4 to 6 grown-up children with fertile marriages are very similar with regard to length of marriage. In about 90% the marriages lasted more than 10 years in all 3 groups, so a difference in this respect cannot reasonably explain the differential viability in the family groups expressed in the different numbers of offspring. The length of the marriage, however, may play a part in the family groups with only one child and especially in the group with none (see the Table). It cannot, however, be the only or even the main factor, because in the former group not less than about three-quarters of the marriages lasted more than 6 years and in more than half the cases more than 10 years. In the family group of generation I with no offspring about half of the marriages lasted more than 6 years and in about one-third more than 10 years. If the last group is divided in two, one with offspring but not of the defined type and another with no offspring at all, in both cases about one-third of the marriages lasted more than 10 years.

The differential viability in the families expressed in the different numbers of defined offspring is difficult to explain. Several factors may be operative and some of them may be genetic in nature. Conditions of nutrition may interact, but the population was homogeneous economically. Venereal diseases were at that time widespread in the population, but we are dealing here with a very stable rural population.

One objection to the calculations above might be that the family group with no offspring with defined

TABLE
PERCENTAGE OF BASIC FAMILIES WITH WIFE
MARRIED DURING FERTILE PERIOD

No. of Families	No. With Fertile Married Children	% Marriages, Wife < 35 Years of Age, With Duration of Marriages			
		> 2 yr.	> 4 yr.	> 6 yr.	> 10 yr.
75	0	57	52	47	36
111	1	86	81	74	55
82	2	99	99	98	85
60	3	100	100	98	93
37	4-6	100	100	100	95

type in generation I should be omitted from the comparison. From a population-genetical point of view, its members are of no interest since they are not represented in the gene pool of the next generation. Even if married, they are of no more interest than their contemporary unmarried adults (without offspring) in the subpopulation. But if we correct for this, we find the same trend, one-third of generation I—families with 3 and more defined offspring—is responsible for as much as 55% of the gene pool of the next generation (II), i.e. the individuals who grow up, marry, and in their turn get offspring.

Thus while there was an increase of about 6% of the married biologically viable part of the subpopulation from the first to the second generation there was also a not inconsiderable shift in the gene pool with regard to origin. (During this period there was an increase in population size in the whole country of about 10%.) This shift must reasonably have a population-genetic effect, whatever may be its cause, genetical, or environmental, or both. In our day we have effective birth control, which restricts the large sibships. Consequently we have to reckon that the mechanism discussed here and observed in the Swedish subpopulation about 150 years ago must be of far less influence today.

A similar conclusion of a shifting of genetic structure in a population was made by Scheidt (1932). He based it on a large genealogical study of a subpopulation of farmers living on the island Finkenwärder in the Elbe from 1629 to 1870, the size of the subpopulation since the middle of the 18th century being more than one thousand.

However, against my own conclusions discussed above, a reservation, and a rather strong one, must be made. The distribution of the first generation agrees in fact very well with a Poisson distribution, with on average 1.7 children per family, $\chi^2 = 1.34$, $0.90 < p < 0.95$, $df = 4$. Such is even the case for that part of the first generation, distributed along the upper line of the Figure and in the Table, which had a duration of marriage of more than 6 years and where the female partner married before the age of 35. They have 1.9 children on average per family, $\chi^2 = 1.18$, $0.80 < p < 0.90$, $df = 4$. A further analysis of the detailed fertility of the different groups of the second generation distributed along the lower line of the Figure, and of the third generation, must be made before any certain conclusions can be made. If the tendency is a random one, an oscillation back again in the following generations may be expected. Material for such an analysis is under collection and investigation.

Summary

That patients with rare recessive diseases are, in a high proportion of cases, offspring of cousin marriages has been known since the early days of clinical genetics. Weinberg deduced a formula for the relation between the prevalence of a recessive disease and the first-cousin marriage rates among the parents of the affected compared with those in the general population. Weinberg's formula has proved to have useful applications, but due attention must be paid to the fact that the cousin marriage rate may be different in different areas, even in the same country.

For measuring the isolate size in human populations, Dahlberg proposed using simply the first-cousin marriage rate in the general population, but for several reasons this has not proved to be successful. The main isolating factor is distance, and a method using this has been elaborated by Wright, and proved to be very successful.

The fertility for the years 1800–1825 has been studied for a Swedish subpopulation of isolate type. It was found that 55% of the second generation came from only 26% of the first generation.

REFERENCES

- Alström, C. H. (1958). First-cousin marriages in Sweden 1750–1844 and a study of the population movement in some Swedish subpopulations from the genetic-statistical viewpoint. *Acta Genetica et Statistica Medica*, **8**, 295–369.
- (1963). Population movement in Sweden from a genetic-statistical point of view. *Acta Psychiatrica Scandinavica*, **39**, 41–46.
- , and Olson, O. (1957). Heredo-retinopathia congenitalis monohybrida recessiva autosomalis. A genetical-statistical study. *Hereditas*, **43**, 1–178.
- Ammann, F., Klein, D., and Böhringer, H. R. (1961). Résultats préliminaires d'une enquête sur la fréquence et la distribution géographique des dégénérescences tapéto-rétiniennes en Suisse (Étude de cinq cantons). *Journal de Génétique Humaine*, **10**, 99–127.
- Bell, Julia (1940). A determination of the consanguinity rate in the general hospital population of England and Wales. *Annals of Eugenics*, **10**, 370–391.
- Dahlberg, G. (1929). Inbreeding in man. *Genetics*, **14**, 421–454.
- (1943). Mathematische Erblichkeitsanalyse von Populationen. *Acta Medica Scandinavica*, Suppl. 148.
- (1948). *Mathematical Methods for Population Genetics*. Karger, Basle.
- Garrod, A. E. (1901). About alkaptonuria. *Lancet*, **2**, 1484–1486.
- (1902). The incidence of alkaptonuria: a study in chemical individuality. *Lancet*, **2**, 1616–1620.
- Hallgren, B. (1959). Retinitis pigmentosa combined with congenital deafness; with vestibulo-cerebellar ataxia and mental abnormality in a proportion of cases. A clinical and genetical-statistical study. *Acta Psychiatrica Scandinavica*, **34**, Suppl. 138.
- Norio, R. (1966). Heredity in the congenital nephrotic syndrome. *Annales Paediatricae Fenniae*, **12**, Suppl. 27.
- Nuutila, A. (1968). Neuropsychiatric and genetic aspects of the dystrophica retinae pigmentosa-dysacusis syndrome. Thesis, Helsinki.
- Scheidt, W. (1932). Niedersächsische Bauern. II. Bevölkerungsbiologie der Elbinsel Finkenwärder vom dreissigjährigen Krieg bis zur Gegenwart. Fischer, Jena.

Sjögren, T. (1931). Die juvenile amaurotische Idiotie. Klinische und erblichkeitsmedizinische Untersuchungen. *Hereditas*, **14**, 197–426.

Weinberg, W. (1920). Methodologische Gesichtspunkte für die

statistische Untersuchung der Vererbung bei Dementia praecox. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, **59**, 39–50.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.