

STANDARD CALCULATIONS FOR EVALUATING A  
BLOOD-GROUP SYSTEM

R. A. FISHER

*Department of Genetics, Cambridge*

Received 18.x.50

## I. THE TYPES OF PROBLEM

THE number of distinct systems of blood-groups of which the serology and genetics are now understood has increased considerably in recent years. These differ greatly in complexity, chiefly owing to the varying numbers of antibodies available. When we have only a single antibody as with Duffy (*Fy*), only two phenotypes are distinguishable; without family evidence, the heterozygote is indistinguishable from one of the homozygotes. There are, of course, six genetically distinguishable types of mating, but in the absence of offspring, or when the origin of possible offspring is in question, only three types of mating are phenotypically distinguishable. In the case of Kell (*K*) two antibodies are known antigenic to two allelomorphic genes, so that three genotypes can be distinguished, and the six possible matings among them. This was formerly the case with Landsteiner and Levine's system *MN*, but more recently Sanger and Race have found a third antibody, anti-*S*, enabling four alleles, or gene combinations, to be recognised, and six phenotypic classes among the ten genotypes. There are thus 21 distinguishable classes among the 55 different matings possible. Four antibodies may be claimed for the *OAB* system, and in the Rhesus system, the most complex so far elucidated, more than six distinct antibodies are known and the number of genotypes and matings, if very rare kinds are included, is very large indeed.

Owing to the high reliability of diagnosis, these systems are all capable of aiding the solution of genetic problems, some of which are of occasional forensic importance. Four types of problem have from time to time been discussed, and the strength of the evidence which each system can supply varies greatly both according to the system used, and to the purpose for which it is required. The four problems I have in mind are as follows:—

- (a) Doubtful identity.  
How frequently will different persons (supposed unrelated) be of the same phenotype?
- (b) Identical twins.  
How frequently will fraternal twins, children of the same father, be of the same phenotype?

## (c) Doubtful parentage.

If an infant is interchanged with another in a maternity hospital, how frequently will the supposed child be of a phenotype compatible with those of its supposed parents?

## (d) Disputed paternity.

How frequently will the child of another man be compatible with falsely accepted paternity?

For any one blood-group system, and indeed for given phenotypes within each system, the four questions will have different answers. In a court of law it is just such particular circumstances that have to be brought into the calculation. We may, however, also wish to assess the power of the system as a whole, or in other words to calculate the proportion of all cases that present themselves, in which testing the parties for a given blood-group system, will fail to make a discrimination.

Since each such test will in practice be applied simultaneously with other tests, and since each test can turn the balance only when the others fail, it is necessary to measure the relative values of different systems, or of applications to different problems, not by the probability itself of failure to discriminate, but, if this is represented by  $P$ , by

$$-\log P = \log (1/P),$$

which will then be additive when different but independent tests are applied together. The smaller  $P$  is the more successful in the discrimination, and the larger is the logarithm by which its power is appropriately measured. We shall use natural logarithms in all such transformations.

## 2. THE MNS SYSTEM

An example which illustrates the main points which arise in other cases is provided by the system of  $M$ ,  $N$  and  $S$ , in which three antibodies are at present available, while a fourth, anti- $s$  may be found at any time.\*

The frequencies of the four postulated gene combinations are given by Race and Sanger as follows, for British samples

$MS$	24.71722	per cent.
$M_s$	28.31308	„
$NS$	8.02080	„
$N_s$	38.94890	„

These combinations are inherited apparently as a set of four allelomorphs, at a single locus; at least recombination has not so far been observed, and the wide departure from proportionality seen above, itself shows that recombination is either absent, or extremely rare. From these frequencies the frequencies of the ten possible genotypes are calculated by finding the four squares for homozygotes,

\* Since this paper was written, the author has been informed on the highest authority that the fourth antibody, anti- $s$ , has now been discovered. 20th April 1951.

and the six products, doubled, for heterozygotes. These are divisible into six phenotypes by the use of the three antibodies *M*, *N* and *S*, and would be divided into nine phenotypes if all four antibodies were available, as shown in the following table.

TABLE 1  
*Genotypes and phenotypes of the system MNS*

Genotypes	Frequencies per cent.	Phenotypes	
		with anti- <i>s</i>	without anti- <i>s</i>
<i>MS/MS</i>	6.1094	<i>MS</i> 6.1094	<i>MS</i> 20.1058
<i>MS/Ms</i>	13.9964	<i>MSs</i> 13.9964	
<i>Ms/Ms</i>	8.0163	<i>Ms</i> 8.0163	<i>Ms</i> 8.0163
<i>MS/NS</i>	3.9650	<i>MNS</i> 3.9650	<i>MNS</i> 27.7611
<i>MS/Ns</i>	19.2542	<i>MNSs</i> 23.7961	
<i>Ms/NS</i>	4.5419	<i>MNs</i> 22.0553	<i>MNs</i> 22.0553
<i>Ms/Ns</i>	22.0553		
<i>NS/NS</i>	0.6433	<i>NS</i> 0.6433	<i>NS</i> 6.8913
<i>NS/Ns</i>	6.2480	<i>NSs</i> 6.2480	
<i>Ns/Ns</i>	15.1702	<i>Ns</i> 15.1702	<i>Ns</i> 15.1702
	100.0000	100.0000	100.0000

Problems of identity may be solved using these values only, for the frequencies of the different matings and of the phenotypes of children from them are not involved. For example, if an individual has been found to be of type *MS* using only three antibodies, the probability of a blood sample from a different individual agreeing with his is 0.201058, while if there were still agreement using four antibodies the probability of a coincidence is reduced to 0.061094. In general, however, since the phenotype of the known subject may be frequent or rare in proportion to these same frequencies, the probability of failing to distinguish two different blood samples (using this blood-group system) is the sum of the squares of the phenotypic frequencies, giving the values shown in table 2.

TABLE 2  
*Values of MNS system, with 3 or 4 antibodies, for determining identity*

	3 antibodies	4 antibodies
Probability of non-discrimination . . . . .	0.200324	0.163548
Power log (1/P) . . . . .	1.60783	1.81066
Power per cent. . . . .	100	112.6

The fourth antibody will, if and when available, increase the power of the system by 12.6 per cent. The power of the system

TABLE 3  
*Frequency of similarity by chance of sibs from different matings*

Frequency	Genotype of mother	Genotype of father											
		MS/MS	MS/Ms	M <sub>s</sub> /M <sub>s</sub>	MS/NS	M <sub>s</sub> /N <sub>s</sub>	MS/NS	M <sub>s</sub> /NS	M <sub>s</sub> /N <sub>s</sub>	NS/NS	NS/N <sub>s</sub>	N <sub>s</sub> /N <sub>s</sub>	
6·1094	MS/MS	100	100	100	50	50	50	50	50	50	100	100	100
13·9964	MS/M <sub>s</sub>	100	62·5	50	50	37·5	37·5	37·5	37·5	25	100	62·5	50
8·0163	M <sub>s</sub> /M <sub>s</sub>	100	50	100	50	50	50	50	50	50	100	50	100
3·9650	MS/NS	50	50	50	37·5	37·5	37·5	37·5	37·5	25	50	50	50
19·2542	M <sub>s</sub> /N <sub>s</sub>	50	37·5	50	37·5	37·5	37·5	25	25	25	50	37·5	50
4·5419	MS/NS	50	37·5	50	37·5	37·5	37·5	25	37·5	25	50	37·5	50
22·0553	M <sub>s</sub> /N <sub>s</sub>	50	25	50	37·5	37·5	37·5	25	25	37·5	50	25	50
0·6433	NS/NS	100	100	100	50	50	50	50	50	50	100	100	100
6·2480	N <sub>s</sub> /N <sub>s</sub>	100	62·5	50	50	37·5	37·5	37·5	37·5	25	100	62·5	50
15·1702	N <sub>s</sub> /N <sub>s</sub>	100	50	100	50	50	50	50	50	50	100	100	100

employed is, as has been explained, additional to the power of any other independent system used in making the same test.

### 3. IDENTITY OF TWINS

In principle the test of identity of twins is the same as that for individual identity ; in this case, however, the two samples tested are known to be from full sibs, so that the sums of squares may be calculated without knowledge of the population frequencies. Thus in different cases we have

Ratio	Sums of Squares	
1 : 1	$(\frac{1}{2})^2 + (\frac{1}{2})^2$	$= \frac{1}{2}$
3 : 1	$(\frac{3}{4})^2 + (\frac{1}{4})^2$	$= \frac{5}{8}$
1 : 2 : 1	$2(\frac{1}{4})^2 + (\frac{1}{2})^2$	$= \frac{3}{8}$
1 : 1 : 1 : 1	$4(\frac{1}{4})^2$	$= \frac{1}{4}$

If the genotypes of the parents are known, we may use the actual fraction appropriate to the particular phenotype observed in the children. For example, from the mating

$$MS/Ms = NS/Ns$$

three-quarters of the offspring will be of phenotype *MNS*, if tested with three antibodies. If both children are of this phenotype we should take this value,  $\frac{3}{4}$ , to be the probability of agreement in type supposing the children are not identical.

In general, as with other tests involving hereditary transmission, it is necessary to consider the results of all different types of mating, and their frequencies. This is conveniently done by making a symmetrical table, with rows for the genotype of the wife, and columns for that of the husband. The frequency of each mating type is the product of the frequencies of the parental genotypes, since all matings between unlike genotypes appear twice in symmetrical situations.

Multiplying the percentages in any row by the frequencies of the corresponding paternal genotype, we obtain average percentages for each genotype of the mother (page 100).

Averaging these again, to include all types of mother, by multiplying by the frequencies of the maternal genotype, we find the proportion of all cases in which discrimination fails as 48.2915 per cent., or

$$\begin{array}{ll} P & 0.482915 \\ \log (1/P) & 0.72792 \end{array}$$

If table 3 is reconstructed to show the increased discrimination which would follow from the use of anti-*s*, we find :—

$$P \quad 0.442769$$

$$\log (1/P) \quad 0.81471$$

an increase in power of 11.9 per cent. ascribable to anti-*s*.

TABLE 4

Genotype of mother	Frequency of genotype	Percentage of non-discrimination
<i>MS/MS</i>	6.1094	75.0918
<i>MS/Ms</i>	13.9964	47.4186
<i>Ms/Ms</i>	8.0163	64.9696
<i>MS/NS</i>	3.9650	43.7730
<i>MS/Ns</i>	19.2542	37.9178
<i>Ms/NS</i>	4.5419	36.0787
<i>Ms/Ns</i>	22.0553	35.7373
<i>NS/MS</i>	0.6433	75.0918
<i>NS/Ns</i>	6.2480	47.4186
<i>Ns/Ns</i>	15.1702	64.9696

It has been remarked that when the genotypes of the parents are known, the probability appropriate to the particular case observed may be obtained at once. If, however, we have only the observation on the children, giving their phenotype, the probability of an accidental coincidence will not be the same for all phenotypes. To calculate the appropriate value we must combine all mating types capable of producing the phenotype observed in proportion to their probability of doing so. For example, if, using three antibodies we find the phenotype *MMS*, this must be of genotype *MS/Ms* or *MS/MS*. The latter may arise from 16 possible matings, and the former from 16 with their 16 reciprocals. The frequencies of these matings may as before be obtained by multiplying the frequencies of the genotypes of their partners. Table 5 illustrates the computational procedure for the chosen example.

TABLE 5

Source of gamete <i>MS</i>		Source of gamete <i>MS</i>				Source of gamete <i>Ms</i>				Sum of products
		<i>MS/MS</i>	<i>MS/Ms</i>	<i>MS/NS</i>	<i>MS/Ns</i>	<i>MS/Ms</i>	<i>Ms/Ms</i>	<i>Ms/NS</i>	<i>Ms/Ns</i>	
Genotype	Frequency	6.1094	6.9982	1.9825	9.6271	13.9964	16.0326	4.5419	22.0553	
<i>MS/MS</i>	6.1094	100	100	50	50	100	100	50	50	62.2425
<i>MS/Ms</i>	6.9982	100	75	50	50	75	50	25	25	42.3758
<i>MS/NS</i>	1.9825	50	50	25	25	50	50	25	25	31.1200
<i>MS/Ns</i>	9.6271	50	50	25	25	50	50	25	25	31.1200
										10.3811
										51.6323
										per cent.

It will be noticed in the columns of genotypes capable of producing *MS* that three have been halved since from these only half the gametes are *MS*; those supplying *Ms*, moreover, have been treated in the same way, save that all have then been doubled to allow for the fact that in the matings in which they are involved *Ms* may be obtained from either parent. The entries in the body of the table are the fractions of the offspring of the phenotype *MMS*. The right-hand column shows the sums of products with the frequencies at the column heads, these in turn are multiplied by the frequencies to the left, added and divided by the general phenotype frequency 20·1058 per cent. The final value for the probability of a coincidence, 0·516323, shows that the phenotype *MMS* is a trifle less evidential than the average. The six phenotypes distinguishable with these antibodies, calculated in this way, give approximately

<i>MS</i>	0·5163	<i>MNS</i>	0·5066	<i>NS</i>	0·4138
<i>Ms</i>	0·4116	<i>MNs</i>	0·4734	<i>Ns</i>	0·4795

4. INTERCHANGE AND PATERNITY

For cases of suspected interchange, where the child in question, if not that of his supposed parents, is presumed to be a child both of a different father and a different mother, we make up a table of all possible matings according to the distinguishable phenotypes of the father and mother, and enter in each cell the sum of the frequencies of all phenotypes which such a mating could produce. Thus in table 6 we have a 6 × 6 table of the matings which the use of three sera can distinguish; for parentage is excluded if the child does not belong to one of these phenotypes. Apart from this, which phenotype it belongs to is a matter of indifference.

TABLE 6

Total frequencies of phenotypes arising from different matings

Frequency	Phenotype of mother	Phenotype of father							
		<i>MS</i>	<i>Ms</i>	<i>MNS</i>	<i>MNs</i>	<i>NS</i>	<i>Ns</i>		
20·1058	<i>MS</i>	28·1221	28·1221	77·9385	77·9385	49·8164	49·8164	57·7245	75·0919
8·0163	<i>Ms</i>	28·1221	8·0163	77·9385	30·0716	49·8164	22·0553	41·3446	64·6732
27·7611	<i>MNS</i>	77·9385	77·9385	100·0000	100·0000	71·8779	71·8779	87·5917	87·5459
22·0553	<i>MNs</i>	77·9385	30·0716	100·0000	45·2418	71·8779	37·2255	66·4206	74·9310
6·8913	<i>NS</i>	49·8164	49·8164	71·8779	71·8779	22·0615	22·0615	54·6835	75·0919
15·1702	<i>Ns</i>	49·8164	22·0553	71·8779	37·2255	22·0615	15·1702	43·7700	64·6732

It will be noticed that the different matings supply evidence of very different strength, and that in 3 cases out of 36 (or 2 out of 21, if complementary matings are counted as one) there is no evidence whatever. Multiplying each column by the paternal frequency and

adding, we have the first column on the right, and these values multiplied by the maternal frequencies give the general average, of 64.2944 per cent., being the frequency of failure to detect an interchange, using three antibodies only.

Had the question been one of paternity only, the values in table 6 should be divided by the sum of the frequencies of the phenotypes which might have been born to the mother without restriction of paternity. These happen to be the values in the third column (headed *MNS*). The second column on the right shows the percentages of wrongly ascribed paternity (for each maternal phenotype) which would not be disproved by the *MNS* system using three antibodies. The average for all maternal phenotypes is 76.0989, showing only a little more than one-quarter successful disproofs. Using four antibodies, on the other hand, the percentage of failures is reduced to 54.3501 for interchange, and 66.8560 for paternity.

We may now compare the power of the system with three antibodies with that in which four antibodies were available, for the four purposes which have been discussed.

*Powers of the tests for different purposes*

	3 antibodies	4 antibodies	Percentage increase
Identity . . . . .	1.60783	1.81066	12.6
Identity of twins . . . . .	0.72792	0.81471	11.9
Interchange . . . . .	0.44170	0.60973	38.0
Paternity . . . . .	0.27312	0.40263	25.7

The order of difficulty of the four types of discrimination discussed is generally the same for all blood-group systems, but this general similarity is compatible with a considerable disproportion of power, as is well illustrated by the comparisons above between the same system with three and with four antibodies. The increase in power which a fourth antibody would give is strikingly large both for interchange and for paternity problems, though only moderate for individual identity, and identity of twins.

## 5. SUMMARY

Four types of genetic or forensic problem, for which blood group evidence is of value, are distinguished; and standard calculations appropriate to each are exhibited, using, in each case, the factor *MNS* as developed by Sanger and Race.

## REFERENCE

- SANGER, R., AND RACE, R. R. 1947. Subdivisions of the *MN* blood groups in man. *Nature*, 160, 505.