

" EPIDEMICS " CAUSED BY A LATE-MANIFESTING GENE. APPLICATION TO SCRAPIE

G. J. DRAPER
Unit of Biometry, University of Oxford

Received 22.ix.62

1. INTRODUCTION

VARIATIONS in the incidence of scrapie in sheep have recently been discussed by Parry (1960 and 1962). The pattern of occurrence of the disease is sometimes suggestive of an epidemic, but Parry's evidence supports the hypothesis that its transmission is in fact largely or wholly controlled by a single autosomal recessive gene.

The object of the present paper is to show that, under certain circumstances, the incidence of a genetically controlled disease can increase dramatically in a comparatively short time, thus suggesting an epidemic caused by an infective agent. The data used for illustrative purposes pertain to scrapie but the phenomenon discussed is a general one which may be expected to occur under the (presumably rare) conditions that :

- (i) the disease concerned is due to a single autosomal recessive gene,
- (ii) it manifests after the age at which reproduction occurs,
- (iii) there is selection in favour of animals carrying the gene.

2. SCRAPIE

A comprehensive survey of various aspects of the disease has been given by Parry (1962), who describes in some detail the selective pressure in favour of animals which (on the present hypothesis) are carrying the scrapie gene (p. 98). It appears that many rams preferred for breeding purposes (*a*) eventually manifest the disease, (*b*) have close relatives affected, or (*c*) have affected progeny. On the genetic hypothesis it is likely that such preferred animals will be homozygous or heterozygous for the postulated scrapie gene. The disease usually manifests between the ages of $2\frac{1}{2}$ and $4\frac{1}{2}$ years, by which time the animal may have several offspring, perhaps a hundred if it is a ram. There is now a good deal of evidence in favour of the suggested genetic mechanism, *e.g.* Parry (1960, 1962), Draper and Parry (1962).

Thus it seems that (i), (ii) and (iii) of Section 1 apply in the case of scrapie. The theory of the following sections is developed in general terms for a situation in which these criteria apply, referring to scrapie when data are required for illustrative purposes.

It should be emphasised that we are not attempting to *prove* that scrapie is controlled by a recessive gene but rather to show that certain

observations made on the spread of the disease are explicable in such terms. Nor is any attempt made to fit theoretical predictions to any observed body of data.

3. VARIATIONS IN GENE-FREQUENCY

The following definitions, notation and formulæ will be required :

The *gene-frequency* of a gene is, as usual, defined as the proportion of available loci occupied by that gene. If the proportions of homozygous recessives, heterozygotes and homozygous dominants are respectively r , h and d ($r+h+d = 1$), the frequency of the recessive gene is

$$p = r + \frac{1}{2}h \quad (1)$$

and that of the dominant is

$$q = \frac{1}{2}h + d = 1 - p \quad (2)$$

The symbols P_i and p_i respectively will be used to denote the frequency of the recessive gene for the breeding males and females of the i th year.

We denote by t_i the gene-frequency for the progeny born in the i th year. Then under the assumption of random mating* it is well known that

$$t_i = \frac{1}{2}(P_i + p_i). \quad (3)$$

The gene-frequency of a group of animals consisting of several sub-groups each with its own gene-frequency is the weighted average of these latter gene-frequencies, the weight appropriate to each sub-group being the proportion of the whole constituted by that sub-group.

The above concepts and formulæ can now be applied to the spread of a disease such as scrapie. As indicated above it is assumed that the males are in some way selected. Their gene-frequency P_i is computed from the proportions of different genotypes used each year. Normally these will not be known (at any rate in the case of scrapie) but for our present purpose we can legitimately invent plausible figures.

The gene-frequencies for the females are computed as follows. (It is assumed that those retained for breeding are representative of their age-group.) We suppose that the females used for breeding in the i th year were born in years $i-1, i-2 \dots (i-j)$ and that they occur in the proportions $a_1, a_2 \dots a_j$ ($a_1 + a_2 \dots + a_j = 1$). Then,

$$p_i = a_1 t_{i-1} + a_2 t_{i-2} \dots + a_j t_{i-j}$$

Using this value of p_i we may compute t_i from formula (3) and proceed recursively computing p and t for any number of successive

* The deviations from this which may occur in practice are immaterial for our purpose, which is merely to provide a general description of the way in which the importation of a "bad" gene can lead to an apparent epidemic.

years. (However, we obviously cannot *compute* the t_i for the earliest years but have to *assume* a gene-frequency for such animals.)

The proportion of homozygous recessive progeny in year i is $P_i p_i$ and this is the proportion of animals in that birth-group * in which the disease should eventually manifest (though in general some animals will be removed by other causes before this happens). An apparent epidemic may occur if the product $P_i p_i$ increases from year to year, particularly if this happens rather suddenly.

Some calculations based on the formulæ given above and on invented data suggested by some case-histories of scrapie flocks are presented in the next section.

4. THE HYPOTHETICAL BREEDING POPULATIONS

Four cases are considered. We take two hypothetical flocks of breeding ewes and two series of rams, and compute for successive

TABLE 1
Age structure of two hypothetical populations of breeding ewes

Age	Percentage of flock in each age-group	
	Flock I (high incidence)	Flock II (low incidence)
1½	40	25
2½	33	22
3½	14	18
4½	9	14
5½	4	10
6½	...	7
7½	...	4

years the proportions of each genotype which would be obtained in the progeny if one of the flocks of ewes were mated with one of the series of rams.

Table 1 gives the age-structures of the two ewe populations. (This age structure is kept constant for each flock by the introduction of young ewes and the gradual removal of older ones.)

Columns 1 and 2 of table 2 show the composition of the two series of rams. These series are constructed so that the (hypothetical) sets of rams used in each year have gene-frequencies of a magnitude which might occur in practice under the circumstances outlined in section 2. It is assumed that three rams are used in each year. Column 1 indicates their genotypes. Column 2 shows the corresponding gene-frequency in each year (the latter is of course the important variable ; on the assumption of random mating the actual types or numbers of animals used are immaterial).

* A “ birth-group ” is defined as the group of animals born into the flock in a given year.

5. SPREAD OF GENE IN THE HYPOTHETICAL POPULATIONS

In table 2A are shown the results which would be obtained by mating each of the two flocks with the first series of rams. Table 2B

TABLE 2

The numbers of homozygous recessive progeny to be expected in two types of flock for the cases when one of two series (A and B) of rams is used

Year	Rams used			Gene frequency of rams	Gene frequency of progeny		Percentage of homozygous recessives among progeny	
	ss	sS	SS		Flock I	Flock II	Flock I	Flock II
<i>Case A</i>								
1	0	0	3	0	0	0	0	0
2	0	3	0	0.50	0.25	0.25	0	0
3	1	2	0	0.67	0.33	0.33	0	0
4	0	3	0	0.50	0.30	0.28	5	3
5	2	1	0	0.83	0.52	0.49	18	12
6	2	1	0	0.83	0.55	0.51	22	16
7	3	0	0	1.00	0.69	0.64	38	28
8	3	0	0	1.00	0.74	0.68	48	36
9	2	1	0	0.83	0.70	0.64	48	38
10	2	1	0	0.83	0.75	0.68	55	44
11	2	1	0	0.83	0.76	0.71	58	48
12	2	1	0	0.83	0.78	0.73	60	52
13	2	1	0	0.83	0.79	0.74	62	55
14	2	1	0	0.83	0.80	0.76	63	57
15	0	0	3	0	0.37	0.35	0	0
<i>Case B</i>								
1	0	0	3	0	0	0	0	0
2	0	1	2	0.17	0.08	0.08	0	0
3	0	2	1	0.33	0.17	0.17	0	0
4	0	1	2	0.17	0.10	0.09	1	0
5	0	2	1	0.33	0.21	0.20	3	2
6	0	1	2	0.17	0.14	0.12	2	1
7	0	2	1	0.33	0.24	0.22	5	4
8	0	1	2	0.17	0.16	0.14	3	2
9	0	2	1	0.33	0.26	0.24	6	5
10	0	1	2	0.17	0.18	0.16	3	3
11	1	1	1	0.50	0.36	0.34	11	9
12	2	1	0	0.83	0.52	0.51	17	15
13	2	1	0	0.83	0.55	0.53	22	19
14	0	2	1	0.33	0.36	0.32	13	10
15	0	1	2	0.17	0.32	0.27	8	6

In the above tables it is to be noted that the crucial factor is the gene-frequency of the set of rams used in each year, and not the individual genotypes.

gives the corresponding results for the second series of rams. The frequency of the recessive gene in each of the two populations of ewes is assumed to be initially zero. In Flock I the proportion of

young ewes is greater ; this explains the differences between the incidence in the two flocks which is apparent both in table 2A and in 2B, for if sires with a high gene-frequency are used the gene will spread less rapidly in Flock II owing to the "damping" effect exerted by the older ewes than in Flock I where "carriers" will soon tend to form a large proportion of the breeding population.

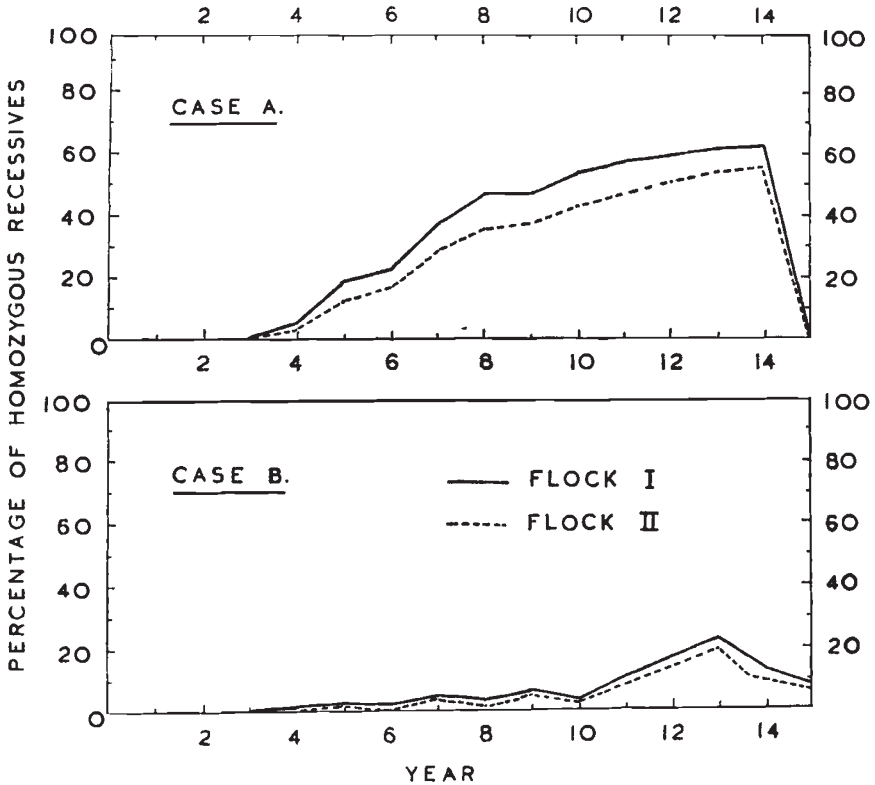


FIG. 1.—Percentage of homozygous recessive progeny in successive birth-groups, for the four hypothetical cases discussed in the text.

Both flocks illustrate the point that if animals homozygous or heterozygous for a recessive gene are used extensively for breeding purposes the pattern of occurrence of a disease depending on the gene may well give the appearance of an epidemic. This is shown particularly for the first series of rams where the birth-group incidence in the progeny rises from zero to 60 per cent. in about ten years. For the second series of rams there is a maximum incidence of about 20 per cent. in each flock, occurring rather suddenly, after a period in which the incidence is around 5 per cent. The importance of this in relation to scrapie is discussed in the next section. These points are illustrated further in figs. 1A and 1B.

6. DISCUSSION

The "model" considered in this paper applies to sheep-breeding practice (in Britain). However, the essence of the conclusions will be unaffected if the crucial element of selection in favour of the gene operates in some other way, *i.e.* if such selection applies to females rather than males, or to both sexes.

The incidence calculated here is that for each *birth-group*; the *annual* incidence, which is more easily calculable in practice, and variations in which are more readily apparent to an observer, will depend on the incidence in several successive birth-groups. This annual incidence will not in general attain such high levels as the birth-group incidence, nor will it show such dramatic variations, for the deaths in any birth-group will be spread over a number of years. The difference between the two will become most marked in cases where manifestation tends to occur late in life. From the practical point of view the importance of this is that variations in the percentage of recessives born may take some years to become apparent.

Also it can be seen from table 2, particularly case B, that the gene-frequency may be quite high for some time while the percentage of recessives in each birth-group remains small. If the gene-frequency of the rams then suddenly increases the percentage of recessives will also do so. It is quite possible in the case of scrapie for cases to be missed or even concealed (for its presence in a flock is regarded with some disapprobation) and thus a low incidence may appear to be zero. But if the birth-group incidence of recessives suddenly increases there will be a sudden increase in the incidence of the disease about three years later which may then be too high to remain unnoticed. It is suggested that it is precisely this sort of circumstance which leads to apparent epidemics of scrapie.

The theory presented is inadequate in at least two respects.

(a) No account has been taken of the fact that since many homozygous recessives are removed by the disease after one or two gestations the gene-frequency in the females of a given birth-group will tend to decrease with time, because the proportion of recessives remaining will decrease. However, since the heterozygotes may also transmit the gene and since our age-structures for the populations of breeding females show a preponderance of young animals (which may contain many non-manifest "affecteds") it seems unlikely that this would alter the general conclusions. Again, by postulating an even "younger" age-structure it might be possible to offset this effect of early removal on the figures presented.

(b) Perhaps a greater defect of the analysis is that it is deterministic and not stochastic, *i.e.* that no account has been taken of the random fluctuations which will occur in practice. However, this is immaterial since the purpose has been merely to illustrate the general nature of the phenomena under consideration. To incorporate all

the conceivable random elements into the “ model ” would probably make the mathematics intractable and would certainly obscure the essential features of the problem.

7. SUMMARY

There is good reason to suppose that scrapie in sheep is transmitted by a single autosomal recessive gene, the effects of which may not be manifest until after the age of reproduction. In addition there appears to be selection in favour of this gene.

It is shown that such a combination of circumstances may lead to a rapid increase in the incidence of the disease, giving the appearance of an epidemic.

The theory is equally applicable to other situations where the same or similar conditions apply, but these are presumably rare.

Acknowledgment.—I am greatly indebted to Mr H. B. Parry of the Nuffield Institute for Medical Research, Oxford for suggesting this investigation and for much helpful discussion and encouragement.

8. REFERENCES

- DRAPER, G. J., AND PARRY, H. B. 1962. Scrapie in sheep : the hereditary component in a high-incidence environment. *Nature (Lond.)*, 195, 670-672.
- PARRY, H. B. 1960. Scrapie : A transmissible hereditary disease of sheep. *Nature (Lond.)*, 195, 441-443.
- PARRY, H. B. 1962. Scrapie : A transmissible and hereditary disease of sheep. *Heredity*, 17, 75-105.

NOTE ADDED IN PROOF

Since this paper was written further doubts have been cast on the suggestion that scrapie is controlled by a recessive gene. The problem remains unresolved; however, it should be pointed out that the results obtained here do not depend on theories about scrapie, though clearly the *relevance* of the paper to scrapie depends on the true aetiology of the disease.