The American Journal of Human Genetics

Vol. 11 No. 2 *June 1959*

Huntington's Chorea in Michigan[1,2]

2. Selection and Mutation

T. EDWARD REED AND JAMES V. NEEL

Department of Human Genetics, University of Michigan Medical School, Ann Arbor, Michigan

INTRODUCTION

THIS PAPER is the second in a series of three describing the results of a comprehensive study of Huntington's chorea in the state of Michigan. The first paper was concerned with the demography and certain aspects of the genetics of this disease (Reed and Chandler, 1958). This presentation will be concerned with selection and mutation in relation to this condition. The third and final communication will be of a clinical nature.

Previous data have suggested that the relative fitness of choreics, compared to normal persons, is in the neighborhood of unity or above (Panse, 1942; S. Reed and Palm, 1951). If this were in fact the case, it would be a rare—not to say unique—situation in human genetics: individuals affected with a severely debilitating disease whose onset is often during the reproductive period nevertheless actually achieving a greater-than-normal fertility. Also, if these studies are correct, it then becomes necessary to explain why Huntington's chorea is a rare disease today, having a frequency in populations of European ancestry of about 4×10^{-5}. This study was undertaken primarily to analyze the population "dynamics" of the gene responsible for this disease in the state of Michigan. It will be shown that not only are individuals heterozygous for this gene at a reproductive disadvantage as compared with normal, but that there are probably at least three and possibly four different ways in which this disadvantage is brought about. The unreliability of sibling controls in studies of genetically determined fertility differentials will be demonstrated for the case of Huntington's chorea. Finally, the rate with which mutation resulting in this phenotype occurs will be calculated.

SELECTION

The essence of this paper is an attempt to estimate the fitness of individuals with Huntington's chorea and to utilize this estimate in several calculations. Estimating fitness is always difficult in man and is especially so for Huntington's chorea. For this reason it is necessary to give considerable attention to the methods which were used.

Received July 5, 1958.

[1] This study was supported in part by contract AT(11-1)-405 of the United States Atomic Energy Commission.

[2] We wish to acknowledge the assistance on clinical aspects of this study given by Dr. Joseph H. Chandler and the excellent field work done by Mrs. Estella Hughes and Mrs. Ruth Davidson.

A. Methodology

The fitness of one class of humans, say population *1* (P_1), relative to that of another, P_2, has usually been defined as $\bar{B}_1/\bar{B}_2$, where the $\bar{B}$'s refer to mean fertility. There are a variety of ways in which the mean fertility may be estimated and a number have been used, some without apparent genetic justification. Reed (1959) has recently discussed the definitions of relative fitness (W) for human populations and proposed the following definition as best: Assume that for P_1 and P_2 we have complete information on the survival and reproduction of a large random sample of liveborn individuals who have been followed from birth to death. Let

x = age in years at last birthday
N_i = original number of newborn liveborn individuals of population i
$B_{i\cdot x}$ = number of livebirths born to the survivors of the N_i individuals during their xth year of life
$\bar{B}_{i\cdot o}$ = mean number of livebirths ever born per newborn liveborn individual of P_i
$B_{i\cdot y}$ = mean number of livebirths ever born per newborn liveborn individual of P_i per year, based on $\bar{B}_{i\cdot o}$ livebirths ever born
$P_{i\cdot x}$ = parental age frequency distribution of the N_i individuals (the proportion of livebirths, out of $\bar{B}_{i\cdot o}$ livebirths, which is born to the survivors of these individuals at age x).

Then,

$$\bar{B}_{i\cdot o} = \frac{1}{N_i}\sum_x B_{i\cdot x}\,,\; P_{i\cdot x} = \frac{B_{i\cdot x}}{\sum_x B_{i\cdot x}}\,,\; B_{i\cdot y} = \frac{1}{N_i}\sum_x \frac{B_{i\cdot x}}{x}\,,$$

and the proposed definition of W for populations *1* and *2*, is

$$W = \frac{B_{1\cdot y}}{B_{2\cdot y}} \tag{1}$$

or its equivalent

$$W = \frac{\bar{B}_{1\cdot o}\sum_x \frac{P_{i\cdot x}}{x}}{\bar{B}_{2\cdot o}\sum_x \frac{P_{2\cdot x}}{x}}. \tag{2}$$

The number of livebirths ever born per newborn is used in order to include all factors affecting genetic fitness: viability, marriage, adult fertility. $B_{i\cdot y}$ is used instead of $\bar{B}_{i\cdot o}$ alone because the latter is in reality a mean number of livebirths *per generation* and we have no assurance that the generation lengths of P_1 and P_2 are equal. Even if equal, the mean number of livebirths ever (to be) born per newborn *per year*, which gives an exact rate of increase of the population due to births, may differ since this is proportional to $\sum_x \frac{P_{i\cdot x}}{x}$. When the parental age distributions are the same, then, as (2) shows, (1) reduces to the usual ratio of means, $\bar{B}_{1\cdot o}/\bar{B}_{2\cdot o}$.

One further point should be noted about what may be called the reference popula-

tion, P_2. Unless otherwise specified, the fitness of a group or genotype should always be related to the general population. The use of other groups, such as normal sibs, for P_2 may be allowed for convenience or necessity but it then becomes necessary to show that P_2 is representative of the general population in its fertility. As will be seen, the non-choreic sibs of the choreic individuals of the present study differ significantly in their completed fertility from the general population of Michigan. Consequently, a direct comparison with sibs would give a misleading estimate. It was therefore necessary to use several different approaches to obtain valid comparisons of choreics (or heterozygotes) with the general population.

Bias due to ascertainment of the kindreds (i.e., obtaining an unrepresentative sample) is believed to be negligible in this study since ascertainment is almost complete, sibships being used as of 1940 and ascertainment having continued until 1956.

In all the calculations to follow, the fitnesses of males and females were calculated separately to test whether there are significant differences. Such a difference has already been found for one dominant trait, namely neurofibromatosis (Crowe, Schull, and Neel, 1956).

In addition to the difficulties in estimating fitness discussed above, there is another peculiar to Huntington's chorea and other dominant traits with delayed onset, namely, the inability, at any specified age, to recognize a certain fraction of heterozygotes. Since the "normal" non-choreic sibs of choreics may still, at any age, in fact be heterozygotes who have failed to manifest the disease, and their children may receive the gene for Huntington's chorea and develop the disease themselves, it is necessary to estimate the number of non-choreic heterozygotes by the use of age-onset curves. This number, added to the number of choreics, gives an estimate of the total number of heterozygotes. It is the fitness of heterozygotes, not choreics, which is of *genetic* importance. The fitness of choreics is, of course, of interest for other reasons. The necessity for this correction, i.e., estimating the number of heterozygotes from the number of trait-bearers, has sometimes been overlooked.

Possible differences in the survival from birth to age five years, between children of choreics and children of non-choreics, were also investigated, since Panse (1942) has found evidence for increased mortality among children of choreics.

B. Results

1. *Sibship comparisons of mean number of children ever born*

For a series of choreics and their non-choreic siblings the conventional ratio of mean numbers of children was calculated, subject to the conditions discussed above. The sibships used are those containing a Michigan choreic living on April 1, 1940, all members of the sibship being used. The status of each member is taken as of the time of death or, if living, the time of last investigation (1954 to 1956). The classification of "choreic" means that either a) a person has been medically diagnosed as having Huntington's chorea (see Part 1 of this study) and/or b) he has been so classified by reliable, non-contradictory lay reports. If only b), he is also a near biological relative (sib, son, etc.) of a medically diagnosed choreic. It is not desirable to restrict "choreic" to medically diagnosed cases since this introduces selection for severity of the disease and increased age of the individual. "Non-choreic" means that there is reliable in-

formation that the individual in question, at the time of investigation, lacked signs of the trait. The three sibships containing "Negro" choreics were excluded as well as all sibships in which information on fertility or choreic status was faulty or lacking. Sibships in which a choreic was ascertained only through an affected descendant, were omitted. The need to estimate the number of non-choreic heterozygotes imposed further restrictions. This was met by calculating the probability C_x that a non-choreic sib of a choreic has the gene for Huntington's chorea but does not show the trait at age x, given that his parent has the gene. Therefore sibships for which one parent could not be considered to be heterozygous for the gene leading to Huntington's chorea (*Hh*) were omitted.

In order to ensure that the reproduction of individuals was nearly, or completely, terminated, only persons living at age 45 or over, or deceased at age 15 or over, were used. Therefore choreics and non-choreics who die relatively young are included in the survey. The age 45 is a compromise between the desire to be certain that reproduction of the still-living is completed and the desire not to select for mildly affected choreics who can live to advanced ages. Age 15 was chosen because no chorea occurred earlier in this sample. Therefore the assumption that the gene has no effect on the fitness of its bearer before age 15 seems plausible, although, of course, this is not certain. This lower age limit enables us to neglect, for this calculation, the many infant and juvenile deaths which, since we cannot determine genotype, contribute no information. After making these omissions, 120 choreic males, 137 choreic females, 97 non-choreic males and 113 non-choreic females were available for study. Five individuals could not be classified either as "choreic" or "non-choreic" from available information. They were omitted from the calculations since preliminary work (Reed, 1957) showed that this omission had a negligible effect.

In order to estimate C_x, we may note that if P_x is the probability that an individual who is *Hh* develops chorea by age x, then

$$C_x = \frac{1 - P_x}{2 - P_x}.$$

P_x may be estimated from the age of onset distribution as was done in Part 1 of this study. Since the values of C_x may be of interest elsewhere, they are presented, with the values for P_x, in Table 1. It may be noted that at age 40, C_x is 0.240 so that the assumption that non-choreic sibs of this age are *hh* (homozygous normal) would lead to gross error. At age 50, C_x becomes 0.063, and at 60, 0.006. C_x at advanced ages is perhaps unreliable because the apparently normal distribution of ages of onset may not, in fact, obtain for extreme deviations from the mean.

In Tables 2 and 3 the distribution by age at death or the age at time of last investigation of the above-described choreic individuals and their non-choreic sibs are presented, together with the number of liveborn children ever born to them. Tables 4 and 5 give the distribution by marital status of the number of liveborn children ever born to these individuals. From the distributions in Tables 2 and 3 the number N of heterozygotes among the non-choreic sibs was calculated from the relation

$$N = \sum_x N_x C_x$$

TABLE 1. VALUES OF P_x AND C_x. (SEE TEXT FOR DEFINITION.) P_x VALUES DERIVED FROM TABLE 15 OF PART 1 OF THIS STUDY, ASSUMING AGE OF ONSET IS NORMALLY DISTRIBUTED WITH MEAN 35.30 YEARS AND STANDARD DEVIATION 9.80 YEARS

Age x	P_x	C_x	Age x	P_x	C_x	Age x	P_x	C_x	Age x	P_x	C_x
15	.019	.495	30	.295	.413	45	.839	.139	60	.994	.006
16	.024	.494	31	.330	.401	46	.862	.121	61	.996	.004
17	.031	.492	32	.367	.388	47	.883	.105	62	.997	.003
18	.038	.490	33	.405	.373	48	.903	.088	63	.998	.002
19	.048	.488	34	.448	.356	49	.919	.075	64	.998	.002
20	.059	.485	35	.488	.339	50	.933	.063	65	.999	.001
21	.072	.481	36	.528	.321	51	.945	.052	66	.999	.001
22	.087	.477	37	.567	.302	52	.955	.043	67	.999	.001
23	.104	.473	38	.610	.281	53	.965	.034	68	1.000	<.0005
24	.125	.467	39	.648	.260	54	.972	.027			
25	.147	.460	40	.684	.240	55	.978	.022			
26	.171	.453	41	.719	.219	56	.983	.017			
27	.198	.445	42	.752	.199	57	.986	.014			
28	.227	.436	43	.785	.177	58	.990	.010			
29	.261	.425	44	.813	.158	59	.992	.008			

where N_x is the number of non-choreics age x at time of death or last investigation. The number B of liveborn children ever born to these N non-choreic heterozygotes is

$$B = \sum_x B'_x C_x$$

where B'_x is the number of livebirths born to non-choreics who were age x at time of death or last investigation, assuming that the fertility of non-choreic heterozygotes is the same as that of all non-choreics. If one makes the more extreme assumption that the fertility of these non-choreic heterozygotes is the same as that of all choreics, the mean is only slightly changed, dropping about 0.005 in the males and increasing about 0.026 in the females. In this situation it seems adequate to make the former assumption. The values of N and B for all groups, together with the numbers of individuals and livebirths and the means for all groups, are given in Table 6.

About ten per cent of the non-choreic sibs in this sample are calculated to be heterozygotes. The number of these heterozygotes plus the number of choreics give the estimated number of heterozygotes in each category. We note that the mean for all non-choreic males, 2.072 $\pm$ 0.260, does not differ from the mean for all non-choreic females, 2.027 $\pm$ 0.202, so that males and females may be pooled to give a better estimate of the fertility of non-choreics. There is, however, a marked and significant difference in the fertility of male choreics and female choreics, the former being 1.850 $\pm$ 0.198, the latter 2.818 $\pm$ 0.233. A large fraction of this difference is explainable by the relatively high proportion of male choreics who fail to marry. In Tables 4 and 5 we see that the proportion of "never married" male choreics is 31/120 = 0.258, while the corresponding proportion for female choreics is 11/137 = 0.080.

TABLE 2. DISTRIBUTION BY AGE AT DEATH OF THE NUMBER OF DECEASED CHOREICS AND DECEASED NON-CHOREIC SIBS, AND THE NUMBER OF LIVEBORN CHILDREN EVER BORN TO THEM. SIBSHIPS CONTAINING A MICHIGAN CHOREIC ON APRIL 1, 1940 AND SELECTED, AS DESCRIBED IN TEXT, FOR ESTIMATING RELATIVE FITNESS. N_x = NUMBER OF INDIVIDUALS AGE x; B'_x = NUMBER OF LIVE-BIRTHS TO INDIVIDUALS AGE x.

Age x	Males				Females				Age x	Males				Females			
	Choreic		Non-Choreic		Choreic		Non-choreic			Choreic		Non-choreic		Choreic		Non-choreic	
	N_x	B'_x	N_x	B'_x	N_x	B'_x	N_x	B'_x		N_x	B'_x	N_x	B'_x	N_x	B'_x	N_x	B'_x
15	0	0	1	0	0	0	0	0	50	6	9	0	0	0	0	1	10
16	0	0	0	0	0	0	0	0	51	3	7	2	4	9	40	1	6
17	0	0	0	0	0	0	2	0	52	4	8	0	0	4	12	0	0
18	0	0	1	0	0	0	0	0	53	5	12	1	3	7	15	1	0
19	0	0	1	0	0	0	0	0	54	2	6	0	0	4	11	0	0
20	0	0	0	0	0	0	0	0	55	3	2	0	0	7	29	1	4
21	0	0	0	0	0	0	1	0	56	3	7	0	0	2	0	0	0
22	1	0	1	0	0	0	1	0	57	2	7	3	14	1	2	0	0
23	0	0	2	0	1	1	1	0	58	1	2	0	0	2	4	3	5
24	0	0	3	1	0	0	2	0	59	1	7	1	4	1	0	1	2
25	0	0	0	0	0	0	0	0	60	1	3	0	0	3	14	1	0
26	1	0	0	0	1	0	1	1	61	5	9	1	0	0	0	1	0
27	1	0	0	0	0	0	2	1	62	0	0	2	10	5	23	1	0
28	0	0	0	0	0	0	1	0	63	3	6	1	0	3	4	0	0
29	0	0	3	2	0	0	1	0	64	5	19	1	2	3	11	0	0
30	1	0	1	0	2	1	3	6	65	2	3	3	15	0	0	0	0
31	2	0	0	0	2	2	1	3	66	0	0	0	0	1	4	0	0
32	0	0	0	0	0	0	1	2	67	3	4	1	1	3	3	0	0
33	1	0	0	0	0	0	1	5	68	0	0	0	0	1	6	1	2
34	0	0	2	4	0	0	1	0	69	1	0	2	7	2	9	1	0
35	1	2	0	0	1	2	0	0	70	2	2	1	0	0	0	2	3
36	1	5	0	0	1	2	0	0	71	0	0	1	0	1	0	0	0
37	5	9	0	0	2	4	0	0	72	0	0	1	2	0	0	0	0
38	1	0	0	0	0	0	0	0	73	2	6	3	5	2	2	1	2
39	1	4	1	0	2	0	3	3	74	0	0	0	0	2	4	0	0
40	2	6	1	0	3	5	2	5	75	1	0	2	4	1	4	1	0
41	2	0	0	0	1	9	0	0	76	0	0	1	7	1	12	1	3
42	2	3	0	0	1	0	0	0	77	1	3	1	1	0	0	1	8
43	1	0	0	0	2	11	1	0	78	1	1	1	0	2	9	0	0
44	3	5	0	0	2	3	1	4	79	0	0	0	0	0	0	1	4
45	6	5	1	0	4	9	0	0	80	1	0	0	0	0	0	0	0
46	1	0	1	0	4	12	2	4	81	0	0	0	0	0	0	0	0
47	3	2	0	0	1	3	0	0	82	0	0	0	0	0	0	1	6
48	1	0	2	11	5	13	0	0	83	0	0	0	0	0	0	0	0
49	2	1	0	0	5	17	0	0	84	0	0	1	1	0	0	0	0
									85	0	0	0	0	0	0	1	4
									Total	97	165	51	98	107	312	50	93

TABLE 3. DISTRIBUTION BY AGE AT TIME OF LAST INVESTIGATION OF THE NUMBER OF LIVING CHOREICS AND LIVING NON-CHOREIC SIBS AND THE NUMBER OF LIVEBORN CHILDREN EVER BORN TO THEM. SIBSHIPS CONTAINING A MICHIGAN CHOREIC ON APRIL 1, 1940 AND SELECTED, AS DESCRIBED IN TEXT, FOR ESTIMATING RELATIVE FITNESS. N_x AND B'_x AS IN TABLE 2.

Age x	Males				Females			
	Choreic		Non-choreic		Choreic		Non-choreic	
	N_x	B'_x	N_x	B'_x	N_x	B'_x	N_x	B'_x
45	0	0	3	7	1	2	1	0
46	0	0	1	1	2	6	7	23
47	0	0	2	0	0	0	0	0
48	2	4	1	1	3	9	1	4
49	2	3	0	0	1	7	3	4
50	0	0	3	2	0	0	3	4
51	3	4	2	5	1	1	2	6
52	0	0	2	2	2	5	1	1
53	1	3	0	0	1	2	2	6
54	2	7	2	2	1	0	0	0
55	0	0	2	6	0	0	2	7
56	1	1	0	0	2	9	2	6
57	0	0	0	0	1	2	2	0
58	2	4	1	4	3	1	3	4
59	0	0	0	0	1	0	2	7
60	1	0	1	0	0	0	1	2
61	1	0	4	15	2	4	1	3
62	2	12	2	10	1	4	1	0
63	1	2	0	0	1	4	1	1
64	1	8	1	0	0	0	5	14
65	1	0	4	8	0	0	0	0
66	1	2	0	0	0	0	2	4
67	0	0	1	1	0	0	2	0
68	0	0	1	0	0	0	5	9
69	0	0	1	0	0	0	1	5
70	1	0	1	2	2	11	1	3
71	0	0	3	6	2	2	1	4
72	0	0	0	0	0	0	1	2
73	0	0	2	5	0	0	2	5
74	0	0	2	6	0	0	2	4
75	0	0	0	0	1	5	0	0
76	0	0	2	10	0	0	3	1
77	1	7	0	0	0	0	0	0
78	0	0	0	0	1	0	1	4
79	0	0	0	0	0	0	0	0
80	0	0	1	2	1	0	0	0
81	0	0	0	0	0	0	1	0
82	0	0	1	8	0	0	1	3
Total	23	57	46	103	30	74	63	136

TABLE 4. DISTRIBUTION OF NUMBER OF LIVEBORN CHILDREN EVER BORN TO THE MALES TABULATED IN TABLES 2 AND 3.
S = SINGLE (NEVER MARRIED); M = MARRIED (EVER MARRIED)

Number of Live-births	Choreic									Non-Choreic								
	Deceased ≥ 15			Living ≥ 45			Total			Deceased ≥ 15			Living ≥ 45			Total		
	S	M	Σ	S	M	Σ	S	M	Σ	S	M	Σ	S	M	Σ	S	M	Σ
0	27	13	40	3	4	7	30	17	47	15	8	23	6	11	17	21	19	40
1	1	12	13	0	3	3	1	15	16	0	7	7	0	3	3	0	10	10
2	0	14	14	0	5	5	0	19	19	0	7	7	0	6	6	0	13	13
3	0	17	17	0	2	2	0	19	19	0	4	4	0	10	10	0	14	14
4	0	4	4	0	2	2	0	6	6	0	5	5	0	2	2	0	7	7
5	0	2	2	0	1	1	0	3	3	0	0	0	0	4	4	0	4	4
6	0	4	4	0	0	0	0	4	4	0	0	0	0	1	1	0	1	1
7	0	2	2	0	1	1	0	3	3	0	1	1	0	1	1	0	2	2
8	0	0	0	0	1	1	0	1	1	0	1	1	0	1	1	0	2	2
9	0	1	1	0	0	0	0	1	1	0	1	1	0	1	1	0	2	2
10	0	0	0	0	1	1	0	1	1	0	1	1	0	0	0	0	1	1
11	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total no. of live-births	1	164	165	0	57	57	1	221	222	0	98	98	0	103	103	0	201	201
Total no. of individuals	28	69	97	3	20	23	31	89	120	15	36	51	6	40	46	21	76	97
Mean no. of live-births	0.036	2.377	1.701	0.000	2.850	2.478	0.032	2.483	1.850	0.000	2.722	1.922	0.000	2.575	2.239	0.000	2.645	2.072
Variance	0.036	3.944	3.941	0.000	7.818	7.715	0.032	4.775	4.700	0.000	8.549	7.554	0.000	5.481	5.519	0.000	6.845	6.547
S.E. of mean	0.036	0.239	0.202	0.000	0.625	0.579	0.032	0.232	0.198	0.000	0.487	0.385	0.000	0.370	0.346	0.000	0.300	0.260

TABLE 5. DISTRIBUTION OF NUMBER OF LIVEBORN CHILDREN EVER BORN TO THE FEMALES TABULATED IN TABLES 2 AND 3.
S = SINGLE (NEVER MARRIED); M = MARRIED (EVER MARRIED)

Number of Live-births	Choreic									Non-Choreic								
	Deceased ≥ 15			Living ≥ 45			Total			Deceased ≥ 15			Living ≥ 45			Total		
	S	M	Σ	S	M	Σ	S	M	Σ	S	M	Σ	S	M	Σ	S	M	Σ
0	9	13	22	0	8	8	9	21	30	11	12	23	9	8	17	20	20	40
1	1	17	18	1	1	2	2	18	20	0	3	3	0	9	9	0	12	12
2	0	22	22	0	7	7	0	29	29	0	8	8	0	14	14	0	22	22
3	0	13	13	0	3	3	0	16	16	0	6	6	0	9	9	0	15	15
4	0	8	8	0	6	6	0	14	14	0	4	4	0	7	7	0	11	11
5	0	2	2	0	2	2	0	4	4	0	2	2	0	2	2	0	4	4
6	0	8	8	0	0	0	0	8	8	0	2	2	0	2	2	0	4	4
7	0	5	5	0	1	1	0	6	6	0	0	0	0	2	2	0	2	2
8	0	4	4	0	1	1	0	5	5	0	1	1	0	1	1	0	2	2
9	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
10	0	1	1	0	0	0	0	1	1	0	1	1	0	0	0	0	1	1
11	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
12	0	2	2	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0
Total no. of live-births	1	311	312	1	73	74	2	384	386	0	93	93	0	136	136	0	229	229
Total no. of individuals	10	97	107	1	29	30	11	126	137	11	39	50	9	54	63	20	93	113
Mean no. of live-births	0.100	3.206	2.916	1.000	2.517	2.467	0.182	3.048	2.818	0.000	2.385	1.860	0.000	2.519	2.159	0.000	2.462	2.027
Variance	0.100	8.145	8.210	0.000	4.687	4.602	0.164	7.390	7.415	0.000	5.717	5.429	0.000	3.764	4.007	0.000	4.534	4.615
S.E. of mean	0.100	0.290	0.277	0.000	0.402	0.392	0.140	0.242	0.233	0.000	0.383	0.330	0.000	0.264	0.252	0.000	0.221	0.202

TABLE 6. NUMBER OF INDIVIDUALS, NUMBER OF LIVEBIRTHS, AND MEAN NUMBER OF LIVEBIRTHS
Data of Tables 2, 3, 4, and 5. Numbers for specified genotypes are estimated. See text

Hh = heterozygote for gene for Huntington's chorea
$(Hh)_n$ = non-choreic Hh individual
hh = homozygous normal individual
B = observed number of liveborn children ever born to the N individuals (age $\geq$ 15 years)
$\bar{B}_o$* = estimated mean number of liveborn children ever born per *newborn* individual
N = number of individuals

Status		Males					Females				
		Chore-ics	Non-chore-ics	$(Hh)_n$	Hh	hh	Chore-ics	Non-chore-ics	$(Hh)_n$	Hh	hh
Deceased	N	97	51	7.839	104.839	43.161	107	50	10.369	117.369	39.631
$\geq$ 15	B	165	98	4.297	169.297	93.703	312	93	11.412	323.412	81.588
	$\bar{B}$	1.701	1.922	0.548	1.615	2.171	2.916	1.860	1.101	2.756	2.059
	S.E.	0.202	0.385	—	—	—	0.277	0.330	—	—	—
Living	N	23	46	1.358	24.358	44.642	30	63	1.884	31.884	61.116
$\geq$ 45	B	57	103	1.979	58.979	101.021	74	136	4.656	78.656	131.344
	$\bar{B}$	2.478	2.239	1.457	2.421	2.263	2.467	2.159	2.471	2.467	2.149
	S.E.	0.579	0.346	—	—	—	0.392	0.252	—	—	—
Total	N	120	97	9.197	129.197	87.803	137	113	12.253	149.253	100.747
	B	222	201	6.276	228.276	194.724	386	229	16.068	402.068	212.932
	$\bar{B}$	1.850	2.072	0.682	1.767	2.218	2.818	2.027	1.311	2.694	2.114
	S.E.	0.198	0.260	—	—	—	0.233	0.202	—	—	—
Total	$\bar{B}_o$	1.388	1.554	—	1.325	1.664	2.170	1.561	—	2.074	1.628

*$\bar{B}_o = \bar{B}l_{15}$ where l_{15} = 0.75 for males and 0.77 for females. These values may be in error by about two per cent. See text.

Mean of males plus females:

	$\bar{B}$	S.E.	$\bar{B}_o$
Non-choreic	2.048	0.162	1.556
hh	2.162	—	1.643

The probability that this difference is due to chance is less than .001. The proportions for non-choreic males and for non-choreic females are 0.216 and 0.177, respectively. This increased marriage rate of female choreics has already been noted in Part 1 of this study. That this sex difference in proportion of married choreics may be only one factor in the difference in mean numbers of children is suggested by the fact that the mean for all married male choreics is 2.483 $\pm$ 0.232 while for all female married choreics it is 3.048 $\pm$ 0.242. These means do not differ significantly but the difference is suggestive. It is worth emphasizing here that these means are the fertilities of individuals who have survived to the age of 15 years. To obtain the mean fertility per (liveborn) newborn individual ($\bar{B}_o$) we must multiply each of the above-calculated means by the proportion of individuals of the specified category who survive to age 15, i.e. by l_{15}. Values of l_{15} are not known precisely since a direct count of deaths in these sibships between birth and age 15 is subject to some error and the division be-

tween heterozygotes and homozygotes cannot be made, while estimating l_{15} from life tables is made difficult by the wide range of birth years and geographical origins of individuals in these sibships. However, it should still be possible to estimate l_{15} for the present study from the life tables in the *United States Life Tables*, 1890, 1901, 1910, and 1901–1910 (U. S. Bureau of the Census, 1921) since these tables are based on data from ten north-eastern states, including Michigan. Since the living members of these sibships in 1940 were about 48 years of age, the l_{15} values for 1890 will be appropriate for individuals of this mean age and are an approximation for all individuals in this sample. For 1890 the l_{15} values are given only for Massachusetts (whites plus non-whites), being 0.706 for males and 0.723 for females. For 1901 and 1910 they are available for the ten states and can be further limited to whites only (native plus foreign-born). The values for these males in 1901 and 1910 are 0.780 and 0.805 respectively, while for females they are 0.807 and 0.831. It therefore seems adequate to consider that for the fertility sample composed of native and foreign-born whites, l_{15} is about 0.75 for males and 0.77 for females. The error in these estimates should not exceed two or three per cent for the general population of Michigan in 1890. Since the earliest observed age of onset was 15 years, it is believed that there is no difference in the l_{15} values of choreics, their non-choreic sibs, and the general population of Michigan. Using these values we note, in Table 6, that now, *measuring from birth*, only choreic females and *Hh* females have a $\bar{B}_o$ greater than two, while individuals in other categories have $\bar{B}_o$ values ranging from 1.325 to 1.664. Since the mean number of children required for exact replacement is two, these results indicate that only these two classes of females, on the average, are replacing themselves in the course of one complete life cycle.

Using the mean numbers of children ever born given in Table 6 we can calculate the relative fitness from ratios of means as is usually done. We use the values actually obtained in our sample since correcting to $\bar{B}_o$ involves multiplying both numerator and denominator by l_{15}. We first calculate two relative fitnesses, $W_{c \cdot n}$, defined as the ratio of the mean for choreics, of specified sex, to the mean for non-choreics of both sexes (since males and females don't differ), and $W_{H \cdot h}$, the ratio of the estimated mean for *Hh* individuals, of specified sex, to the estimated mean for *hh* individuals of both sexes. [Because of the many comparisons of relative fitness to be made, the following terminology is adopted. The subscripts of W indicate the two groups whose fitnesses are being compared, the first subscript indicating the numerator and the second the denominator. The subscripts are c (choreic), n (non-choreic sib), H (estimated Hh), h (estimated *hh* sibs), and p (general population). It should be noted at the outset that $W_{H \cdot p}$ (for the mean of males and females) is the relative fitness of genetic importance. The other relative fitnesses are interesting for other reasons or were used in obtaining the mean $W_{H \cdot p}$.] We find the following values, uncorrected for parental age distribution:

	$W_{c \cdot n} \pm$ S.E.	$W_{H \cdot h} \pm$ S.E.
Males	0.903 ± 0.120	0.817 ± 0.110
Females	1.376 ± 0.157	1.246 ± 0.142
Mean of males and females	1.140 ± 0.117	1.032 ± 0.105

The bias of these estimates is negligible compared with the standard error since the standard errors of the mean number of livebirths are about 0.1 of the mean (Cochran, 1953). The standard error of a ratio of means requires knowledge of the correlation between pairs of observations used in estimating the numerator and denominator. This correlation is assumed to be zero here since, in the present data, there is no definite mechanism operating to produce a correlation between the fertility of a choreic, taken at random, with the fertility of a non-choreic, taken at random. A positive correlation would reduce the above standard errors. In testing the significance of the W's the differences between means were used. Only $W_{c \cdot n}$ for females is significantly different from unity ($P < 0.01$). For both sexes $W_{H \cdot h}$ is about ten per cent less than that of $W_{c \cdot n}$ so that the relative fitness of heterozygotes (relative to sibs, not to the general population) is less than a simple comparison of affected with normal sibs would indicate. A fairly striking difference between males and females shown by both $W_{c \cdot n}$ and $W_{H \cdot h}$ is apparent, in each case significant at the 0.01 level.

As shown previously (Eq. 2) the definition of W should include the parental age distribution if these distributions differ in the two groups being studied. In Table 7 the distribution is given of individuals by exact age at the birth of a liveborn child. In Table 8 the distribution of these parental age groups, by five year intervals, is given, together with the parental age (P_x) for males and females in Michigan in 1935 (the earliest year for these data). Using expected values from these Michigan P_x values, the significance of departures of the observed values from the state (testing within sexes) was calculated. As is shown, only female choreics differ significantly ($P < 0.01$), having an excess of births at high ages. The parental age distributions, within sexes, cannot be shown to differ between choreics and non-choreics. Since the $W_{c \cdot n}$ and $W_{H \cdot h}$ values obtained must later be used in estimating fitness relative to the general population, the values based on female choreics must be weighted by

$$\sum_x \frac{P_{1 \cdot x}}{x} \Big/ \sum_x \frac{P_{2 \cdot x}}{x}$$

(see Eq. 2) to correct for the different parental age distribution, where subscript *1* refers to female choreics and *2* to the females of Michigan. From Table 8, using the mid-ages, 17.5, 22.5, etc., this ratio is calculated to be 0.969. The corrected values of $W_{c \cdot n}$ and $W_{H \cdot h}$ are then:

	$W_{c \cdot n}$	$W_{H \cdot h}$
Females	1.333	1.207
Mean of males and females	1.118	1.012

2. *Comparison of fertility of choreic and non-choreic sibs with 1940 census fertility data*

The 1940 Census (April 1, 1940) obtained information on the number of (liveborn) children ever born to white women of specified ages in Michigan. This permits direct comparisons to be made with the fertility of choreics and their non-choreic sibs who were living in Michigan on April 1, 1940. These data are presented in Table 9. It is important to note several characteristics of these data. They give mean number of livebirths per woman *surviving* to the specified age and therefore are not the desired mean nunber of livebirths per newborn individual. This latter statistic is not avail-

TABLE 7. DISTRIBUTION OF INDIVIDUALS IN TABLES 2 AND 3 BY AGE AT BIRTH OF LIVEBORN CHILD. ONLY INDIVIDUALS ALL OF WHOSE AGES AT BIRTH ARE KNOWN

Deceased $\geq$ 15 years plus living $\geq$ 45 years

x = Age at birth of child

N = Total number of individuals

B = Total number of children born to N individuals

Age x	Males		Females		Age x	Males		Females	
	Choreic	Non-choreic	Choreic	Non-choreic		Choreic	Non-choreic	Choreic	Non-choreic
15	0	0	0	1	35	7	4	10	4
16	0	0	1	1	36	4	7	10	3
17	1	2	4	2	37	7	5	6	3
18	1	0	7	5	38	3	2	4	2
19	1	4	9	7	39	3	4	7	4
20	1	2	19	10	40	6	2	9	1
21	5	3	23	5	41	3	2	2	0
22	7	5	17	9	42	2	3	2	0
23	8	2	13	12	43	2	2	3	0
24	8	9	20	9	44	3	2	3	0
25	17	6	16	7	45	1	1	1	0
26	11	10	18	7	46	2	1	0	0
27	15	3	14	7	47	2	2	0	0
28	13	6	14	10	48	3	0	1	0
29	8	10	10	9	49	0	2	0	0
30	9	5	14	4	50	1	0	0	0
31	9	8	4	11	51	2	1	0	0
32	9	4	13	4	53	0	0	0	1
33	12	6	10	6	54	1	0	0	0
34	8	7	9	3	56	1	0	0	0
					N	65	41	86	51
					B	196	132	293	147
					$\bar{x}$	31.158	31.068	27.843	26.823
					S.E.$\bar{x}$	0.539	0.649	0.401	0.515

able for any representative American population. This limitation does not, however, invalidate direct comparison between the means of the census and those of non-choreic males and non-choreic females living at the time of the census because the viability of these normal sibs should not differ appreciably from that of the general population. There are no census data for males but it is obvious that, in the absence of migration, the mean number of livebirths ever to be born to a newborn male must be very nearly equal to that of a newborn female. For ages before the end of the reproductive period the mean age differentials between spouses will make the age-specific means differ, but for ages of 40 or over this difference should be negligible. The non-choreic sibs may therefore be directly compared to the general population of females of the state in terms of fertility of surviving individuals, but this is not the

TABLE 8. COMPARISON OF OBSERVED PARENTAL AGE DISTRIBUTIONS WITH THE DISTRIBUTION IN MICHIGAN IN 1935. DATA OF TABLE 7 AND OF U. S. BUREAU OF THE CENSUS (VITAL STATISTICS—SPECIAL REPORTS)

Age Group	Number of Livebirths								Proportion of Livebirths, Michigan, 1935	
	Males				Females					
	Choreics		Non-choreics		Choreics		Non-choreics		Males	Females
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.		
15–19	3	2.35	6	1.58	21	32.82	16	16.46	.012	.112
20–24	29	35.67	21	24.02	92	92.88	45	46.60	.182	.317
25–29	64	56.25	35	37.88	72	77.06	40	38.66	.287	.263
30–34	47	44.69	30	30.10	50	50.40	28	25.28	.228	.172
35–39	24	30.18	22	20.33	37	28.42	16	14.26	.154	.097
40–44	16	16.66	11	11.22	19	10.26	1	5.15	.085	.035
45+	13	10.00	7	6.73	2	0.88	1	0.44	.051	.003
Total	196	195.80	132	131.86	293	292.72	147	146.85	.999	.999
χ_5^2*	4.53		0.58		16.22		3.07		—	—
P	>0.3		>0.9		<0.01		>0.5		—	—

* Ages 15–24 pooled for males, 40+ for females.

case for the choreics. In this latter case a bias arises as a consequence of using choreics who have reached a specified age. This produces a selection for choreics who are, on the average, less severely affected and may, therefore, differ in fertility from the mean for all choreics. One might expect that they would be more fertile because their chorea is milder. Because of this bias it was not thought worthwhile to make the additional corrections for a) the choreics who may have been selected, wholly or in part, because of having had children, and b) for non-choreics who are in fact heterozygotes and should be added to the choreics to obtain the total heterozygotes. Because of these biases it is very likely that the means for choreics are really an upper limit of the true mean for all choreics and for all heterozygotes. Since 22 individuals living in Michigan in 1940 and non-choreic at that time later developed Huntington's chorea, they are included in Table 9 in the category "Pre-choreic." The sum of choreics and "pre-choreics" should be a good approximation to the actual number of heterozygotes in 1940. Conversely, by not classifying the "pre-choreics" as non-choreics, the latter should be a good approximation to the homozygous normals.

Table 9 shows that, except for choreic females and the few "pre-choreics," individuals in almost every age group are less fertile, but usually not significantly so, than white Michigan females of the corresponding age groups. To obtain an approximation to the number of children ever born and increase the numbers of individuals in a specified category, the fertility of individuals age 40 and over, classified by whether ever married or not, was analyzed. These data are given in Table 10. It is seen that (total) male choreics are much less fertile than the state population of the same age distribution as these choreics (mean 1.867 $\pm$ 0.289; expected mean 2.836). This difference is significant at the .001 level. For total male choreics plus male pre-

choreics the difference is still significant (2.047 ± 0.317; expected 2.824) at the .02 level. The difference between observed and expected for male choreics (with or without pre-choreics) who were ever married is not significant but in each case the fertility of choreics or choreics plus pre-choreics is less. The only other significant difference for separate sexes is that between total non-choreic females and the expected (2.255 ± 0.292; expected 2.902), significant at the .02 level. However, the mean for total non-choreic males, 2.224 ± 0.373, differs from the expected 2.933, at $P = .07$, and if one makes a similar comparison for total non-choreic males age 30 or over, the difference is significant ($P = .02$), the mean again being lower than expected. In each of these cases where a significant difference is found using total individuals, the mean based on "ever married" individuals is also lower than the corresponding expected mean, but not significantly so. The mean fertilities of non-choreics do not differ significantly between sexes and therefore the fertilities of males plus females, for "ever married" and for total individuals, were calculated. For the former category the observed mean is 2.589 ± 0.249, expected 3.086, $P < .05$, and for the latter 2.240 ± 0.232, expected 2.917, $P < .01$. Since the mean number of livebirths ever born per non-choreic age 40 or more ($\bar{B}_{n\cdot 40}$) is less than that of the general population of Michigan females, ($\bar{B}_{m\cdot 40}$) it is very likely that the mean for newborn non-choreics ($\bar{B}_{n\cdot o}$) is also significantly less than the corresponding population mean ($\bar{B}_{m\cdot o}$). It is therefore inadmissible to consider the non-choreics as representative of the general population in any estimation of the relative fitness of choreics or heterozygotes. If one can assume that $\bar{B}_{n\cdot 40}/\bar{B}_{n\cdot o}$ is very near to $\bar{B}_{m\cdot 40}/\bar{B}_{m\cdot o}$ (this appears reasonable, see Reed [1959]), one can estimate the fitness ($W_{n\cdot p}$) of non-choreic sibs relative to that of the general population. We estimate this as

$$W_{n\cdot p} \doteq \frac{\bar{B}_{n\cdot 40}}{\bar{B}_{m\cdot 40}} = \frac{2.240 \pm 0.232}{2.917} = 0.768 \pm 0.080,$$

the mean value for males plus females.

3. *Estimate of the fitness of* Hh *individuals relative to the general population*

Using normal sibs and the 1940 census: If we multiply the estimates of $W_{c\cdot n}$ (corrected for parental age distribution) of the preceding section by $W_{n\cdot p}$ we will obtain estimates of fitness of choreics relative to the general population itself. We may call these estimates $W_{c\cdot p}$. To obtain an estimate of the fitness of heterozygotes relative to the general population ($W_{H\cdot p}$) we need to multiply the fitness of heterozygotes relative to non-choreic sibs ($W_{H\cdot n}$) by $W_{n\cdot p}$. The values of $W_{H\cdot n}$ obtained from Table 6, and including the previous parental age correction, are 0.863, 1.315, and 1.089 for males, females, and their mean, respectively. These estimates are:

	$W_{c\cdot p} \pm$ S.E.	$W_{H\cdot p} \pm$ S.E.
Males	0.694 ± 0.117	0.663 ± 0.107
Females	1.024 ± 0.164	0.979 ± 0.148
Mean of males and females	0.859 ± 0.129	0.821 ± 0.116

The variance of the product of two means, $\bar{x}\bar{y}$, is, in part, a function of the correlation between x and y, increasing with positive correlation and vice versa. There appears

TABLE 9. FERTILITY OF CHOREICS, THEIR NON-CHOREIC SIBS, AND WHITE FEMALES OF THE GENERAL POPULATION. ALL INDIVIDUALS LIVING IN MICHIGAN ON APRIL 1, 1940. DATA AS OF APRIL 1, 1940. SEE TEXT CONCERNING BIASES IN THE MEANS OF CHOREICS RELATIVE TO OTHER MEANS

N = No. of individuals
B = No. of liveborn children (Livebirths in "1940" are counted ¼)

Age	Ever Married	Item	Males				Females				White Females, Michigan, 1940**
			Choreic	Pre-choreic*	Choreic + pre-choreic	Nonchoreic	Choreic	Pre-choreic*	Choreic + pre-choreic	Non-choreic	
15-29	No	N	3	2	5	2	3	2	5	3	—‡
		B	0	0	0	0	1	0	1	0	—
		$\bar{B}$	0	0	0	0	0.333	0	0.200	0	—
	Yes	N	2	2	4	10	3	1	4	10	277,020
		B	0	1.25	1.25	6.50	2	1	3	10	342,716
		$\bar{B}$	0	0.625	0.313	0.650	0.667	1.000	0.750	1.000	1.237
		S.E.	—	—	—	0.256	—	—	—	0.258	—
	Total	N	5	4	9	12	6	3	9	13	622,240
		B	0	1.25	1.25	6.50	3	1	4	10	342,716
		$\bar{B}$	0	0.313	0.139	0.542	0.500	0.333	0.444	0.769	0.551
		S.E.	—	—	0.139	0.224	0.224	—	0.176	0.231	—
30-39	No	N	2	0	2	2	0	0	0	1	—
		B	0	0	0	0	0	0	0	0	—
		$\bar{B}$	0	—	0	0	—	—	—	0	—
	Yes	N	11	4	15	11	17	3	20	20	307,140
		B	21	5	26	16	50	5	55	35	668,540
		$\bar{B}$	1.909	1.250	1.733	1.455	2.941	1.667	2.750	1.750	2.177
		S.E.	0.368	—	0.316	0.312	0.539	—	0.481	0.428	—
	Total	N	13	4	17	13	17	3	20	21	343,380
		B	21	5	26	16	50	5	55	35	668,540
		$\bar{B}$	1.615	1.250	1.529	1.231	2.941	1.667	2.750	1.667	1.947
		S.E.	0.368	—	0.311	0.303	0.539	—	0.481	0.416	—
40-49	No	N	5	0	5	1	2	0	2	2	—
		B	1	0	1	0	0	0	0	0	—
		$\bar{B}$	0.200	—	0.200	0	0	—	0	0	—

40–49	Yes	*N*	17	3	20	10	29	5	34	15	269,180
		B	37	14.25	51.25	27	90	22	112	33	766,772
		$\bar{B}$	2.177	4.75	2.563	2.700	3.103	4.400	3.294	2.200	2.849
		S.E.	0.472	2.529	0.552	0.817	0.531	1.123	0.482	0.536	—
	Total	*N*	22	3	25	11	31	5	36	17	289,720
		B	38	14.25	52.25	27	90	22	112	33	766,772
		$\bar{B}$	1.727	4.750	2.090	2.455	2.903	4.400	3.111	1.941	2.647
		S.E.	0.407	2.529	0.481	0.779	0.515	1.123	0.473	0.504	—
50+	No	*N*	6	0	6	5	4	0	4	6	—
		B	0	0	0	0	0	0	0	0	—
		$\bar{B}$	0	—	0	0	0	—	0	0	—
		S.E.	0	—	0	—	—	—	—	0	—
	Yes	*N*	17	0	17	33	42	0	42	32	373,520
		B	46	0	46	82	135	0	135	91	1,186,644
		$\bar{B}$	2.706	—	2.706	2.485	3.214	—	3.214	2.844	3.177
		S.E.	0.452	—	0.452	0.469	0.475	—	0.475	0.376	—
	Total	*N*	23	0	23	38	46	0	46	38	393,420
		B	46	0	46	82	135	0	135	91	1,186,644
		$\bar{B}$	2.000	—	2.000	2.158	2.935	—	2.935	2.395	3.016
		S.E.	0.417	—	0.417	0.429	0.454	—	0.454	0.359	—
Total‡‡	No	*N*	16	2	18	10	9	2	11	12	—
		B	1	0	1	0	1	0	1	0	—
		$\bar{B}$	0.063	0	0.056	0	0.111	0	0.091	0	—
		S.E.	0.063	—	0.056	0	0.111	—	0.091	0	—
	Yes	*N*	47	9	56	64	91	9	100	77	1,226,860
		B	104	20.50	124.50	131.50	277	28	305	169	2,964,672
		$\bar{B}$	2.213	2.278	2.223	2.055	3.044	3.111	3.050	2.195	2.416
		S.E.	0.260	0.999	0.266	0.292	0.296	0.824	0.278	0.229	—
	Total	*N*	63	11	74	74	100	11	111	89	1,648,760
		B	105	20.50	125.50	131.50	278	28	306	169	2,964,672
		$\bar{B}$	1.667	1.864	1.696	1.777	2.780	2.545	2.757	1.899	1.798
		S.E.	0.227	0.855	0.299	0.265	0.282	0.767	0.264	0.214	—

* Non-choreic on April 1, 1940, choreic later.
** U. S. Census. Data on total white women, except ever married and not reporting on children.
‡ Single women were not asked about children.
‡‡ Means for total choreics and total non-choreic sibs are not comparable with the total state population since the age distributions are very different.

TABLE 10. FERTILITY OF CHOREICS AND THEIR NON-CHOREIC SIBS AND THE EXPECTED FERTILITY ($\bar{B}_m$) OF WHITE FEMALES IN MICHIGAN WITH SIMILAR AGE DISTRIBUTION. ONLY INDIVIDUALS AGE 40 YEARS OR OVER. DATA AND SYMBOLS OF TABLE 9

Ever Married	Item	Males				Females			
		Choreic	Pre-choreic	Choreic + pre-choreic	Non-choreic	Choreic	Pre-choreic	Choreic + pre-choreic	Non-choreic
No	N	11	0	11	6	6	0	6	8
	B	1	0	1	0	0	0	0	0
	$\bar{B}$	0.091	—	0.091	0	0	—	0	0
Yes	N	34	3	37	43	71	5	76	47
	B	83	14.25	97.25	109	225	22	247	124
	$\bar{B}$	2.441	4.750	2.628	2.535	3.169	4.400	3.250	2.638
	$S.E.$	0.325	2.528	0.359	0.403	0.353	1.123	0.338	0.308
	$\bar{B}_m$	3.013	2.849	3.000	3.101	3.043	2.849	3.030	3.072
Total	N	45	3	48	49	77	5	82	55
	B	84	14.25	98.25	109	225	22	247	124
	$\bar{B}$	1.867*	4.750	2.047*	2.224	2.922	4.400	3.012	2.255*
	$S.E.$	0.289	2.528	0.317	0.373	0.339	1.123	0.327	0.292
	$\bar{B}_m$	2.836	2.647	2.824	2.933	2.867	2.647	2.854	2.902

$\bar{B}$ for non-choreic males plus non-choreic females:

	Observed	Expected	P
Ever married	2.589 ± 0.249	3.086	$<.05$
Total	2.240 ± 0.232	2.917	$<.01$

* = different from expected at 5 per cent level of significance.

to be no obvious cause for positive correlation between $W_{c\cdot n}$ and $W_{n\cdot p}$ or between $W_{H\cdot n}$ and $W_{n\cdot p}$. There is some reason, however, to believe that any correlation might be negative, since the mean fertility of non-choreics occurs in the denominator of $W_{c\cdot n}$ and $W_{H\cdot n}$ and in the numerator of $W_{n\cdot p}$. If this were so, the above standard errors, which assume no correlation, would be reduced slightly. Using the standard errors as calculated (which are based on the standard errors for $W_{c\cdot n}$ and $W_{H\cdot h}$, believed to be slightly too large) $W_{c\cdot p}$ and $W_{H\cdot p}$ for males are seen to be significantly different from unity, indicating that male choreics and male *Hh* are less fit than the general population. Females, and the mean of males and females, however, do not differ significantly. However, since the value for the mean $W_{H\cdot p}$, 0.821, (which is *the* relative fitness of genetic importance) is based half on the fitness of male *Hh* which do differ significantly ($P < .01$) from unity, it really is different from unity. This comparison therefore indicates that heterozygotes are less fit genetically than the general population, having a relative fitness of about 0.82 of normal.

Using estimate of $\bar{B}_o$ obtained from Cohort Fertility: The preceding estimate of the mean $W_{H\cdot p}$ was partly based on the mean fertilities of living normal sibs and the 1940 census fertility data because a direct comparison of choreics or heterozygotes with census data is not valid. An estimate of $W_{H\cdot p}$ not dependent on the non-choreic sibs is desirable, however. The fertility data assembled by Whelpton (1954) can be used for such an estimate. Using all available United States census data on fertility

and mortality, Whelpton (1954) has calculated, among other quantities, the mean number of liveborn children ever born ($\bar{B}_s$) to U. S. native-born white women, of certain specified cohorts, who *survive* to specified age x. ("Cohort" here is all of the above-specified liveborn females who are born in a particular year, say 1890). The mean number of liveborn children per original member of the cohort ($\bar{B}_o$) is not given here, nor is it available elsewhere. $\bar{B}_s$ and $\bar{B}_o$ differ appreciably for high x, $\bar{B}_s$ exceeding $\bar{B}_o$, because $\bar{B}_s$ is based only on women who survive to, say, 40 years, neglecting those who die below this age. If one knew the age-specific cumulative fertility of women dying at age x, it would be possible to calculate $\bar{B}_o$ exactly. This fertility is unknown but, as a first approximation, one can assume it is the same as that of women living at age x. It is possible to think of reasons why it should be lower; it is also possible to find reasons why it may be higher. Our interest is in the cohort of 1890 since this approximates the present data well. By multiplying the number of women dying at age x (obtained by subtraction from Table C, Whelpton, 1954) by the cumulative fertility at age x (mean value for age x in Table A, *ibid.*), and summing these products from birth to age 47, the number of births to women dying before age 47, B_d, is obtained. B_d plus the total births from women reaching age 47 or over gives the total births from the original cohort. For the cohort of 1890 these data are not given before age 30 so that it is necessary to estimate these early births from data on the cohort of 1900. (The authors are indebted to P. K. Whelpton, Scripps Foundation for Research in Population Problems, for suggesting this general procedure. The calculations are our own.) If the cohort of 1890 numbered 10,000 liveborn females, it may be calculated that 536 children are born to women dying under 30, 1,538 to women dying between 30 and 46 inclusive, and 17,753 to women surviving to 47 or over. The total of 19,827 children is equivalent to 1.9827 livebirths ever born per original member of the cohort. The accuracy of this estimate is somewhat uncertain but it is clear that the true value is appreciably greater than the minimum estimate of 1.7753, the value which would result if women dying under 47 had no children. An absolute maximum, clearly too high, is given by assuming that all women reaching age 15 survive through age 47. This maximum is 2.340 children ever born. The estimate of 1.9827 is below replacement value, but this is believed to have been characteristic for several native white American cohorts of this period (P. K. Whelpton, personal communication). Since this estimate is for native-born white women of all the United States, we may approximate the corresponding value for Michigan white (native-born plus foreign-born) women by multiplying 1.9827 by the ratio, for women age 45–49 years in 1940, of a) number of livebirths ever born per Michigan white women to b) number of livebirths ever born per U. S. native white woman. This ratio is $2.730/2.602 = 1.049$. The estimate for a Michigan cohort of 1890 is then 2.080, slightly over replacement. This estimate of $W_{H \cdot p}$ requires an estimate of $\bar{B}_o$ for all Hh individuals. From Table 6 and from the weighting factor of 0.969 to correct for the parental age distribution of female choreics, this is $[1/2]\ [1.325 + 0.969 \times 2.074] = 1.668$. The mean $W_{H \cdot p}$ is then $1.668/2.080 = 0.802$.

Using the estimate of $\bar{B}_o$ for the general population derived by Reed (1959): Reed (1959) calculated $\bar{B}_o$ for the normal sibs of sporadic (parents normal) propositi of individuals affected with multiple neurofibromatosis, using published and unpublished

TABLE 11. PROPORTION OF LIVEBORN CHILDREN OF CHOREICS AND OF NON-CHOREIC SIBS DYING UNDER FIVE YEARS OF AGE. ONLY CHILDREN BORN FIVE OR MORE YEARS BEFORE THE TIME OF LAST INVESTIGATION TO WHITE, MARRIED PARENTS

	Male Parent		Female Parent	
	Choreic	Non-choreic	Choreic	Non-choreic
Total number of children	314	266	529	360
Number dying under 5 years	15	10	40	17
Proportion dying	0.048 ± 0.012	0.038 ± 0.012	0.076 ± 0.011	0.047 ± 0.011

Proportion dying under five, non-choreic parents, males plus females: 0.043 ± 0.008

data of Crowe *et al.* (1956). These propositi were ascertained in Michigan between 1934 and 1953. Because their parents and ancestors were normal these sibs are believed to be representative of the general population. A comparison of their completed fertility with data of the 1950 census confirms this belief; no significant differences were found. For 107 such individuals who had lived to age 40 years or over, or who died at any age, $\bar{B}_o$ was found to be 2.037 ± 0.240 livebirths ever born. Using this value, the mean $W_{H \cdot p}$ is estimated to be 1.668/2.037 = 0.819. The relatively large standard error of $\bar{B}_o$ makes this estimate less reliable than the preceding two but it is useful as a check.

The "best" estimate of $W_{H \cdot p}$: Since the estimates vary only from 0.802 to 0.821, the range being less than the smallest standard error, the question of which is "best" is rather academic. It seems reasonable to use the mean of the three estimates as the "best" estimate. This is 0.81, to two decimals.

4. *Infant mortality among the children of choreics and non-choreic sibs*

Panse (1942) reported an increase in infant mortality among children of choreics, (28.64 ± 1.8 per cent dying in the first 10 years of life for children born 1880–1899) relative to that in children of non-choreic sibs of choreics (22.64 ± 2.08 per cent in the first 10 years). He does not distinguish the sex of the choreic parent. This possible mode of selection was investigated in the present study by considering mortality between birth (livebirth) and the fifth birthday. These data are presented in Table 11. The proportion of deaths among children of male non-choreics does not differ significantly from that of female non-choreics so we may pool these groups to obtain an estimate for non-choreics, in general, of the proportion of liveborn children dying before their fifth birthday, 0.043 ± 0.008. The proportion for male choreics does not differ significantly, being 0.048 ± 0.012, but that of females does, being 0.076 ± 0.011 ($\chi_1^2 = 4.958$, $P = .025$). If this difference between male and female choreics is real it is not surprising since the mother is probably more important to an infant's survival than is the father. These data indicate that the real fitness of married female choreics should be decreased by about 0.033 and that of all choreics by about 0.015.

MUTATION

Almost all rare dominant traits in man with high penetrance, when intensively studied, give examples of mutation. Dentinogenesis imperfecta (Witkop, 1957) is

perhaps the only such trait which, after a large-scale investigation, yielded no evidence for mutation. It is therefore appropriate to look for mutation in Huntington's chorea even though, because of its high relative fitness and late onset, good examples are expected to be rare.

The possibility of very late onset of the symptoms of chorea in individuals who are *Hh* is a serious handicap to any attempt to demonstrate mutation. In fact, we consider it impossible, with the present diagnostic facilities, to pick out individual families in which mutation has occurred. At the same time, if one studies a large number of choreic families it may be possible to demonstrate with considerable reliability that mutation is occurring in some of the families. The latest onset of choreiform movements among 204 choreic individuals in the present study was 65 years, this onset age being found in two individuals. The next highest age was 54 years, again occurring in two individuals. These ages indicate the difficulty of proving mutation in specific cases in the face of the alternative of late onset in a parent. An instructive example of this difficulty can be given. Through correspondence we learned of a patient, male, age 44 years, with Huntington's chorea whose parents were reported normal at ages 77 years (father) and 68 years (mother). The patient's wife, the family physician, and the superintendent of the nearby mental institution to which the patient was committed, all stated that the parents and all other relatives were normal. Dr. J. H. Chandler, then of this Department, visited the family (Kindred #4655) (which lives outside of Michigan), and made neurological examinations on the patient, his brother (age 41 years), and both parents. Typical Huntington's chorea was found in both the patient and his brother. Definite, but mild, symptoms of this disease were also observed in the father (e.g. flexion-extension movements at fingers and wrists, jerking movements of shoulders, torsions of the trunk, fleeting universal flexion of lower extremities). The mother appeared entirely normal. Since the family (whose members were intelligent) did not recognize any abnormality in the father, his age at onset is unknown. The important point is that without careful examination this family would have been thought to offer good evidence for mutation, considering the ages of the parents. Other workers have also noted very late onset, 70 years or later (e.g., Entres, 1921; Bell, 1934; Brothers, 1949).

Although we do not believe that specific instances of mutation in Huntington's chorea can be demonstrated, we have recorded, as a matter of interest, the number of instances where a) both parents of a *single* case of Huntington's chorea reached the age of 60 or over and b) were reported to be normal. The kindreds containing parents meeting both requirements are therefore possible examples of mutation. This procedure, however, cannot be used to obtain an estimate of the mutation rate because, obviously, mutation can occur in young parents as well as old and these young parents are excluded. In 196 apparently separate kindreds (groups of biologically related individuals) of Huntington's chorea, one or more of whose choreic members (medically diagnosed) lived in Michigan at some time, there was sufficient information (although usually second-hand) to classify the parents. In eight of the 196 kindreds both were at least 60 and had only one choreic child. Only four of these eight kindreds could be thoroughly investigated and of these four, the father in one was found to have committed suicide at age 62, raising the question of his being in the early stages

of the disease. In the kindred presenting the strongest case for mutation (Kindred ✱4455) the propositus, who was the only case of Huntington's chorea in the kindred, died at age 38 years. Autopsy findings, including brain sections (reviewed by Dr. J. H. Chandler), were typical of the disease. Four sibs, ranging in age from twenty to forty-four years, were reported to be normal. The parents, both age 65 years, were visited by a trained fieldworker, familiar with early and late stages of the disease, and were found to be normal. The parents had no knowledge of the disease in their own ancestors. These facts definitely suggest that mutation has occurred within the last several generations of this kindred but, in view of the fact that this kindred is one of 196 kindreds, we may have merely selected for late onset in the heterozygous ancestors of the patient, all of whom died before showing symptoms of the disease. It is therefore not justifiable to conclude that this example demonstrates mutation. The possibility of diagnostic error, when there is only one choreic in a kindred, is further reason to be cautious.

If, as seems probable, mutation from the normal allele to an allele causing Huntington's chorea does occur, it is obvious that an upper limit of the frequency of such mutation, μ [mutations/locus (loci, *if* there are several producing the phenotype of Huntington's chorea)/generation], can be found. If we classify each individual in Michigan on April 1, 1940 who was known to be *Hh*, (either being choreic at that time or developing chorea at a later date) according to whether or not one of his parents appeared to be *Hh*, being classified *Hh* because of being choreic or, if not choreic, having two or more choreic children and/or a collateral or antecedent choreic relative, we find the following: Of 231 *Hh* individuals known, 206 had an *Hh* parent, 7 were not known to have an *Hh* parent, and 18 could not be classified, usually because they had no known relatives in Michigan. A maximum estimate of the proportion of *Hh* individuals lacking an *Hh* parent, p, is clearly given by $25/231 = 0.108$. The frequency of *Hh* individuals in Michigan, f, was estimated in the first part of this study as 1.01×10^{-4}. Therefore, whether or not the population is in equilibrium with respect to origin and loss of *H* genes, $\mu_{max} = pf/2 = 5.4 \times 10^{-6}$. This is clearly a maximum estimate and probably is several times too high.

No useful estimate of the minimum mutation rate can be derived from the present data. The fact that of 231 *Hh* individuals only 7 are reported to have had normal parents (neither known to be *Hh*) would seem reasonable even if we knew that there were no mutation. If, in fact, there were no mutation this small proportion of *Hh* parents might be expected not to show chorea because of failure to reach the age of onset. Examination of the ages at death or time of last investigation of these seven pairs of parents is not very informative. These ages are, giving the father's ages first: (75, 21), (70, 25), (63, 40), (48, 71), (80, 55), (62, 70), and (69, 87). The P_x and C_x statistics, previously derived from the age of onset distribution, if reliable at advanced ages, could be used to calculate the probability that both members of some of these pairs of parents are *hh* and, therefore, demonstrate mutation. Unfortunately, such reliability cannot be assumed, but, even if it existed, the problem of sampling error would be very difficult. For these reasons no estimate of the minimum mutation rate is attempted.

If the population is in genetic equilibrium with respect to Huntington's chorea,

which is quite possibly not the case, we may estimate μ from $\mu = (1 - W)f/2$, where W is the relative fitness estimated previously, 0.81. Using this value of W, μ is estimated to be 9.6×10^{-6}, almost twice the previously calculated μ_{max}. Since the 95 per cent confidence interval of μ_{max}, based on the estimation of p, does not include 9.6×10^{-6}, and the true value of μ is believed to be at least several times less than μ_{max}, there is reasonably good evidence that either the estimated value of W is in error or that the assumption of genetic equilibrium is incorrect. f is used in both estimates and is therefore not implicated here. It is quite possible, of course, as a result of sampling error or some unknown bias, that W is in error. For example, if the true μ is 2×10^{-6}, the observed f is correct, and genetic equilibrium exists, then W should be 0.96. On the other hand, the assumption of genetic equilibrium for Huntington's chorea seems inherently dubious. The possibility of recent change in the social and psychological characteristics of our society, which can affect the reproduction of *Hh* individuals, is obvious. Although no decision between these two alternatives, or a combination of them, can be made, lack of genetic equilibrium seems most likely to be the explanation. It may be noted in passing that these data demonstrate the uncertainty of indirect estimates of μ when W is near unity since small changes in W may produce large changes in $1 - W$.

In summary, several examples of possible mutation were observed but, because of the possibility of late onset of chorea definite instances of mutation could not be demonstrated. A direct *maximum* estimate of the mutation rate was calculated: 5.4×10^{-6}. If mutation occurs it probably does so at an appreciably lower rate than this.

DISCUSSION

The general difficulties in estimating relative fitness for specific traits, as well as the additional difficulties in the particular case of Huntington's chorea, have been discussed at some length by Reed (1959). The present study has shown how these difficulties are made acute when the relative fitness is high, around 0.8 (mean $W_{H \cdot p}$) in this case, instead of being under 0.5 as is the situation for most genetic traits for which estimates of W have been obtained. Biases which are completely negligible when W is near zero may be very important when W approaches one, thus making the definition of W more difficult than has usually been the case. Examples of such biases are a) age differences between affected and normal sibs (since fertility is age-dependent) when incompleted fertilities are used, b) fertility differences between the normal sibs of the affected and the fertility of the general population, and c) differences in the parental age distribution ($P_{i \cdot x}$ as defined on page 108). The definition of W proposed in equations (1) and (2) is designed to eliminate a) and c). Bias a) is obvious but bias b) and bias c) apparently have not been recognized previous to the findings of this study. The demonstration that b) occurs made it necessary to make all estimates with reference to the general population of Michigan ($W_{H \cdot p}$) and not with respect to the unaffected sibs ($W_{c \cdot n}$) as is usually done. The increasingly available age-specific census data on number of children ever born makes this procedure practical in a number of countries.

Bias in W as a consequence of its usual definition of the ratio of mean numbers of

children uncorrected for parental age distributions, has not usually been recognized. Reed (1959) discussed this and derived equations (1) and (2) to correct for it. The finding that the parental age distribution of female choreics differs significantly from the females of Michigan provided an example of this type of bias.

The estimate of 0.81 for Huntington's chorea and the estimate of 0.78 for multiple polyposis of the colon, (Reed and Neel, 1955), are among the highest relative fitnesses determined for rare dominant or sex-linked recessive traits, clearly a consequence of the late age of onset in these traits. It is probable that there are rare genetic traits, for example dentinogenesis imperfecta, with higher W's but such estimates do not seem to have been made.

It is not known why the normal non-choreic sibs were less fertile than the females of Michigan. Reed and Chandler (1958) found no significant differences between these sibs and the general population with regard to occupation and marital status and the general impression gained was that there were no remarkable differences. An obvious possibility, for which we have no evidence, is that in kindreds having a number of affected persons some of the normal members, or their spouses, limit reproduction because of the fear of developing the disease or of transmitting it to their children. Since the fertility of "ever married" non-choreics as well as total non-choreics is decreased, another possibility is that non-choreics marry later, on the average, than the general population. Unfortunately, we do not have data on age at marriage for non-choreics. Although the mechanism causing this reduction in fertility is unknown, it is an interesting example of "gene action" in man. No other published examples of decreased fertility of normal sibs of individuals with genetic diseases are known to the authors. Data on the fertility of non-choreic sibs given by Panse (1942), S. C. Reed and Palm (1951), and Kishimoto *et al.* (in press) do not permit a comparison with census fertility data since distributions by age and year are not presented. It is worth noting that a tabulation by Reed (1959) of the original data of Crowe, Schull, and Neel (1956) on neurofibromatosis, obtained in a partial survey in the state of Michigan, revealed no indication of a difference in fertility between normal sibs of sporadic propositi and the female population of Michigan. For the normal sibs of familial (one parent affected) propositi, however, the mean fertility was appreciably, but not significantly, lower than the females of Michigan. Since most normal sibs of choreics have an affected parent, the question seems raised whether the presence of an affected parent, or other ancestor, depresses the fertility of unaffected individuals. Further data are clearly required.

The reduced fitness of *Hh* individuals (i.e., the mean for males and females) may have several explanations. The proportion of choreics who marry could not be shown by Reed and Chandler (1958) to differ significantly from the general population but the proportion may, in fact, be lower because male choreics appear to have a lower marriage rate. A more definite factor is the termination of reproduction because the disease is sufficiently advanced to require either institutionalization or other segregation. To estimate the magnitude of this factor, use can be made of the "Relative Reproductive Span" (*RRS*) (Reed and Neel, 1955), assuming (for this calculation only) that the only cause of reduced fitness is institutionalization of some choreics before the end of their potential reproductive period. The age at time of first insti-

tutionalization from Table 16 of Reed and Chandler (1958) was used, equating this age to the termination of actual reproduction, and the distribution of parental ages for Michigan in 1935 (Reed and Neel, 1955) was also used. Under these assumptions W is about 0.93. If, as is probable, fertility is reduced before this age, W would be less. Since the estimate of 0.81 doesn't differ significantly from 0.93 this factor alone might explain all of the reduction in fitness, but this does not seem likely.

The significant difference in fitness observed between male and female choreics, and also *Hh* individuals, is similar to that noted by Panse (1942). He found (see his Table 20) that male choreics had a mean of 3.13 children while female choreics had a mean of 3.56, the ratio of these means being 0.88. The significance of this difference cannot be determined from his data but it probably does not quite reach significance. Kishimoto *et al.* (in press), using data from Japanese pedigrees, report a mean fertility of 3.53 for male choreics and 3.88 for female choreics. These differences are very probably not significant. Other data on fertility, by sex, of choreics do not appear to be published and, in fact, the only other dominant trait similarly analyzed seems to be neurofibromatosis (Crowe, Schull, and Neel, 1956). [The calculation of Vogel (1957) on retinoblastoma, which found no difference between males and females in fertility, employed unilaterally affected adults of varying ages.] Using the method advocated by Krooth (1955), the relative fitness of affected males was found to be 0.413, while that of affected females was 0.748. The significance of these differences was not determined. Calculations by Reed (1959), using the original data of Crowe *et al.*, of the mean completed fertility gave values of 0.451 ± 0.208 livebirths for affected males and 1.292 ± 0.383 livebirths for affected females. These means do not quite differ significantly ($P = .07$) but the difference is suggestive. Although, except for the present study, we lack strong evidence for the reality of these differences, it is interesting that in each study the males appear less fertile. In each case (except for the data of Kishimoto *et al.*, which do not mention the marriage frequency) this seems to be in part, a result of a lower marriage rate among affected males. A reasonable explanation for much, if not all, of the decrease in the proportion of male choreics who marry, relative to that of female choreics, is given by the higher age at time of first marriage of males compared to females. If m_x is the proportion, out of all first marriages in the general population, which occur at age x, and P_x, as defined previously, is the probability that an *Hh* individual will develop Huntington's chorea by age x, then the probability that an *Hh* person marries at age x is approximately proportional to $m_x(1 - P_x)$. $M = \sum_x m_x(1 - P_x)$ should be approximately proportional to the probability that an *Hh* person ever marries. Data to estimate m_x are not available for Michigan before 1950, but data for the neighboring state of Wisconsin for 1922 (when many of the choreics in our study were marrying) should be suitable. $M_{\sigma^{\prime}}$ is found to be 0.752 and M_{φ} is 0.834; their ratio is 0.902. The mean age at marriage is 28.18 years for males and 24.62 years for females; these are slightly high because the data are for all marriages, not first marriages. (Mean age at first marriage for Michigan in 1950 was 24.88 years for males and 21.89 years for females.) From Tables 4 and 5 the observed proportions of male choreics ever marrying is seen to be 0.742, for female choreics 0.920, giving a ratio of 0.807. Thus, the fact that, on the average, a male *Hh* individual is less likely to be non-choreic at the time of his

(actual or potential) marriage than is a female *Hh* individual, seems capable of explaining much of the observed difference. This mechanism is a good example of interrelationship of genetic and social factors in determining fertility. Other possible explanations might be considered but they appear to lack reasonable support. Another example of such interrelationships is the increased mortality found among the children of female choreics but not of male choreics. This increased mortality decreases the relative fitness for all choreics by only about 1.5 per cent from the values estimated on the assumption of no difference in mortality, but is a good example of selection operating at a stage of the life cycle not usually examined.

There appear to have been only three studies, in addition to the present one, in which numerical estimates of the relative fertility of choreics have been obtained. Of these, only Panse (1942) has fertility data from a complete, or nearly complete, survey of a large population, in this case the Rhineland of Germany. In his Table 20, Panse presents data on the number of children born to 457 choreics and 505 non-choreic sibs, all age 30 or over and apparently ascertained through their parents. One can calculate from his data that the mean number of children of choreics is 3.344, that of non-choreics 2.837, giving an estimate of $W_{c \cdot n}$ of 1.18. This is an excellent agreement with the estimate of 1.14 found in the present study (uncorrected for parental age). There is a bias which will increase his value of $W_{c \cdot n}$ although it probably is not of major importance. This bias is a consequence of his minimum age of 30 since, for the age range 30–45, say, choreics will have a higher mean age than their non-choreic sibs because the probability of ascertaining choreics increases with age but the probability for ascertainment of non-choreic sibs does not. This greater mean age in turn produces a greater mean fertility. Panse did not estimate $W_{H \cdot h}$ but it seems probable that such an estimate again would not differ appreciably from our own. He did not compare the choreics' fertility with census fertility data but instead had his own data on 219 normal, more distant, relatives (often first cousins) of choreics, age 30 and over, whose fertility was known. But since the mean fertility of these relatives is 1.973 children, it is difficult to believe that they adequately represent the general population. In fact, since they are close relatives of choreics, their lower fertility raises the question of whether the same mechanism is responsible for their lowered fertility as for that of the normal sibs in the present study. It is therefore not known whether the non-choreic sibs differ from the general population of the Rhineland.

S. C. Reed and Palm (1951) presented data on the fertility of an unstated number of choreics and their non-choreic sibs in the state of Minnesota, U. S. A. They reported a mean number of 6.07 $\pm$ 0.9 children born to choreics and 3.33 $\pm$ 0.5 born to non-choreic sibs. These figures are from the data of Palm (1953) who indicates that they are based on 29 choreic and 49 non-choreic individuals, respectively, who were married and had at least one livebirth. Palm (1953) also reports a significant (at the five per cent level) difference in the mean fertility of 34 choreics and 60 non-choreics unselected for having had children, the respective means being 5.15 $\pm$ 0.87 and 2.72 $\pm$ 0.46. The age distributions are not reported nor is the method of ascertaining these individuals. It is not stated that choreic parents who were ascertained only through their children were excluded from the calculations. The ratio 5.15/2.72 =

1.89 differs markedly from the $W_{c \cdot n}$ ratios found by Panse and in the present study. It is possible that, among other factors, such as age difference between choreics and non-choreics, the fact that they ascertained only a small fraction of the Huntington's chorea kindreds in Minnesota accounts for their high value for $W_{c \cdot n}$. When ascertainment is quite incomplete the probability of ascertaining a kindred is approximately proportional to the number of choreics (or the number known to official sources). As a result there is selection for large kindreds, and consequently, fertile choreics. S. C. Reed and Palm (1951) also discuss a single large kindred, finding that the total number of descendants of a choreic man is four times that of his normal brother. It is not demonstrated that this kindred is representative of choreic kindreds in Minnesota and the mode of ascertainment of the kindred does not exclude bias for large size and hence, fertile choreics. Their data clearly show that in some kindreds choreics are very fertile, but it does not follow that the mean fertility for the state is as high as their results suggest. In our data, for example, there are many kindreds whose choreics are indeed very fertile; there are also others in which choreics are conspicuously less fertile than their normal sibs, usually because of early onset, institutionalization, and death. S. C. Reed and Palm (1951) and Palm (1953) also report that a tabulation of published pedigrees also shows greater fertility (but not significantly so) for choreics than for non-choreic sibs. Since, except for those of Panse (1942), none of these pedigrees are from complete surveys, the above criticisms concerning age distribution, method of ascertainment, and bias for large size also apply here. Because of these biases we do not believe that the data reported by S. C. Reed and by Palm are adequate for estimating the relative fitness of choreics, not to mention that of heterozygotes.

The third study is that of Kishimoto *et al.* (in press) and Kishimoto (personal communication) in Japan, the only data not on Caucasian populations. Fifty-five Japanese kindreds, published and unpublished, were used for the fertility calculations. The mean age of the choreics, living and dead, was about 51 years. No restrictions as to age or ascertainment are mentioned. They found that 58 male choreics had 3.53 children, 56 female choreics had 3.88 children, and 23 sibs (from Aichi Prefecture only, mean age 51 years) had 5.71 children. Using these figures they estimate relative fitness for choreics as 0.649. However, Kishimoto (personal communication) reports that if one compares only choreics and non-choreic sibs studied in his survey of Aichi Prefecture (numbers and ages not specified), the relative fertility of choreics is 0.85. An alternative estimate of the fertility of the general population of Japan can be obtained from the 10 per cent sample tabulation of the 1950 census. The mean number of children ever born per Japanese woman age 50–54 years in 1950 was 4.719. Valid comparisons are not possible since the age and year distributions of choreics are not given and their data are too heterogeneous and subject to bias for large size to yield a valid estimate of W. It does appear, however, that in Japan choreics (and therefore heterozygotes) are also less fertile than the general population, especially considering that, because of biases, the reported choreic fertilities are probably too high. It seems quite possible that the relative fitness of heterozygotes in Japan may be less than that in populations of European ancestry and it is tempting to correlate this with the higher fertility of the general population in Japan. Kishimoto

et al. also reported a frequency estimate for Huntington's chorea, in Aichi Prefecture, of 3.8×10^{-6}, about one-tenth that found in the U. S. A. and in Europe. For a fixed mutation rate and genetic equilibrium, lower relative fitness should produce a lower frequency. Whether this relationship explains these differences between Japan and other countries must remain speculative.

Estimation of mutation rates in dominant traits with late onset like Huntington's chorea and multiple polyposis of the colon (Reed and Neel, 1955) is of necessity indirect and inferential. Nevertheless we believe that the procedure followed in the present study provides a firm upper limit of the mutation rate. We consider it almost certain that mutation does occur since Huntington's chorea is now, and in the past has been, a rare disease, implying a relative fitness less than one. Recurrent mutation seems to be the only way in which the present frequencies can be maintained. The inability to specify a minimum rate of mutation was a consequence of the definite possibility of non-penetrance as an alternative to mutation.

Kishimoto *et al.* (in press) estimate a mutation rate for Huntington's chorea by the indirect method, using their estimate of W and frequency. Their estimate is 6.7×10^{-7}. As mentioned above, this estimate of W is subject to several biases and the calculation is not made in terms of heterozygotes, but this estimate may give the order of magnitude. Several authors, e.g. Panse (1942), mention the possibility of mutation but make no estimates. Our estimate of 5.4×10^{-6} is one of the lowest human mutation estimates made, in keeping with the high estimate for relative fertility of about 0.8. It seems unlikely that accurate or even minimum estimates will be possible until diagnostic techniques improve.

SUMMARY

This paper is the second of three papers based on data from a survey of all known cases of Huntington's chorea (H.c.) in the lower peninsula of the state of Michigan, U. S. A. The first paper (Reed and Chandler, 1958) was concerned with certain demographic and genetic aspects while the third will consider clinical features of the disease.

For the study of natural selection all sibships containing one or more individuals with H.c. living in the lower peninsula of Michigan on April 1, 1940 were considered. From these sibships choreic individuals and non-choreic sibs meeting the following requirements were chosen: 1) classification as "choreic" or "non-choreic" is adequate, 2) ascertainment was independent of fertility, 3) the individual was living at age 45 years or over, or was deceased at age 15 years or over, and 4) one parent of these sibships was very probably heterozygous for the gene for H.c. Two hundred fifty-seven choreics and 210 non-choreic sibs met these requirements. Requirement 3) ensured that only essentially completed fertilities were being measured. Requirement 4) was necessary for the estimation from the age of onset distribution, of the proportion of non-choreic sibs who were in fact unrecognized heterozygotes for the gene for H.c. This proportion was found to be about 10 per cent.

The relative fitness of choreics, when compared to their non-choreic sibs, is about 1.12 ± 0.12, while the relative fitness of heterozygotes (i.e., overt choreics plus individuals heterozygous for the gene but thus far without chorea), compared to homozy-

gous normal sibs, is about 1.01 ± 0.11. The fertility of male choreics (or heterozygotes) was significantly less than that of female choreics (or heterozygotes).

A significant difference between the fertility of the non-choreic sibs and the general population was found. This comparison utilized non-choreic sibs living on April 1, 1940, who were 40 or more years old, and fertility data of the 1940 U. S. census on females having the same age distribution. The relative fertility of the non-choreic sibs is about 0.77 ± 0.08 of the latter. The estimate of the relative fitness of individuals heterozygous for the H.c. gene, *compared to the general population*, is about 0.81.

For the study of mutation, 196 kindreds (groups of biologically related individuals) were available. No specific instances of mutation in H.c. could be demonstrated because non-penetrance could not be definitely excluded. It is noteworthy, however, that in only eight of these kindreds were both parents of a single case of H.c. 1) 60 years of age, or older, and 2) normal. Because mutation occurs in almost all well-studied rare dominant traits in man and also because the relative fitness of H.c. heterozygotes is less than that of the general population, we believe that mutation occurs in H.c. From consideration of the age and diagnostic status of the parents of the heterozygous individuals living in Michigan on April 1, 1940, a direct estimate of the *upper limit* of the mutation rate was obtained. This estimate is 5×10^{-6} mutations per locus (loci, *if* there are several) per generation. An estimate of the lower limit could not be made.

These and other findings are discussed. The reasons for the higher estimates of relative fitness reported in other studies of H.c. are considered.

ACKNOWLEDGMENTS

The authors wish to express their thanks to Dr. Friedrich Vogel and Dr. William J. Schull for their comments.

REFERENCES

BELL, JULIA. 1934. *The Treasury of Human Inheritance.* Vol. 4. Nervous diseases and muscular dystrophies. Part 1. Huntington's chorea. Cambridge: Cambridge University Press.

BROTHERS, C. R. D. 1949. The history and incidence of Huntington's chorea in Tasmania. *Proc. R. Australas. Coll. Physicians* 4: 48–50.

COCHRAN, W. G. 1953. *Sampling Techniques.* New York: Wiley and Sons.

CROWE, F. W., SCHULL, W. J., AND NEEL, J. V. 1956. *A Clinical, Pathological, and Genetic Study of Multiple Neurofibromatosis.* Sprinfigeld: C. C Thomas.

ENTRES, J. L. 1921. Über Huntingtonsche Chorea. *Zschr. ges. Neur. Psychiat.* 73: 541–551.

KISHIMOTO, K., NAKAMURA, M., AND SOTOKAWA, Y. On population genetics of Huntington's chorea in Japan. *Proc. First Internat. Cong. Neurol. Sc.*, Brussels, 1957. (In press.)

KROOTH, R. S. 1955. The use of the fertilities of affected individuals and their unaffected sibs in the estimation of fitness. *Am. J. Human Genet.* 7(4): 325–359.

PALM, J. D. 1953. *Detection of the Gene for Huntington's Chorea.* Ph.D. Thesis. University of Minnesota.

PANSE, F. 1942. *Die Erbchorea.* Leipzig: Thieme.

REED, S. C. AND PALM, J. D. 1951. Social fitness versus reproductive fitness. *Science* 113: 294–296.

REED, T. E. 1957. A genetic survey of Huntington's chorea in the state of Michigan. Proc. Symp. Twin Research Pop. Genetics, Tokyo, 1956. Supplement, *Jap. J. Human Genet.* 2: 48–53.

REED, T. E. 1959. The definition of relative fitness of individuals with specific genetic traits. *Am. J.*

REED, T. E. AND CHANDLER, J. H. 1958. Huntington's chorea in Michigan. 1. Demography and genetics. *Am. J. Human Genet.* 10: 201–225.

REED, T. E. AND NEEL, J. V. 1955. A genetic study of multiple polyposis of the colon (with an appendix deriving a method of estimating relative fitness). *Am. J. Human Genet.* 7: 236–263.

VOGEL, F. 1957. Neue Untersuchungen zur Genetik des Retinoblastoms (Glioma retinae). *Zschr. menschl. Vererb.* 34: 205–236.

WHELPTON, P. K. 1954. *Cohort Fertility.* Princeton: Princeton University Press.

WITKOP, C. J. 1957. Hereditary defects in enamel and dentin. *Acta Genet. et Stat. Med.* 7: 236–239.