Amer J Hum Genet 25:388–396, 1973

Equilibrium Frequencies in X-linked Recessive Disease

SUSAN M. HOLLOWAY[1] AND CHARLES SMITH[1]

Haldane's [1] formula for the equilibrium frequency of rare X-linked recessive diseases maintained by mutation can be extended to cover a wide variety of situations in genetic counseling, antenatal diagnosis, and eugenic consequences of different medical practices. Some of these have been considered already [2–4] but the studies have dealt with the effects of changing one factor at a time. In this paper, formulas are developed so that the net effect of any combined set of factors can be considered, in order to study the balance among different factors and their combined equilibria. The initial sections of the paper introduce the form of the procedure for simple cases; then the methods are generalized to take into account several variables concurrently. Finally, the use of the formulas is illustrated and discussed.

SIMPLE EQUILIBRIA

To introduce the methods and notation used (see Appendix), simple cases dealing with survival rates of affected males and reproductive practices of affected males and carrier females are considered. Let M and F be the equilibrium frequencies at birth of affected males and carrier females, respectively. Let the relative reproductive fitness (number of offspring born relative to the number of offspring born from normal individuals) of affected males and of carrier females be m and f, respectively. Note that this deals with the actual or achieved fertility, rather than with the potential fertility, genetic or otherwise.

Affected males are either the result of a new mutation or are the offspring of carrier females (half of whose sons are affected). Then, if μ is the mutation rate per gamete,

$$\mathrm{M} = \mu + Ff/2. \tag{1}$$

Similarly, carrier females can arise by a new mutation (in either gamete), from carrier mothers (half their daughters are carriers), or from affected males (all daughters are carriers), so

$$F = 2\mu + Ff/2 + Mm. \tag{2}$$

The equilibrium frequencies can then be found, in terms of the mutation rate (μ), by substituting equation (1) in (2) and solving for F, giving

Received December 5, 1972; revised January 29, 1973.

[1] University Department of Human Genetics, Western General Hospital, Edinburgh, EH4 2HU, Scotland.

$$F = \frac{\mu(2 + m)}{[1 - (f/2)(1 + m)]}, \tag{3}$$

and similarly for M.

These formulas can deal with changes in the fitness of affected males and carrier females as a result of genetic counseling or through improved treatment or survival of affected males. For example, if the fitness of affected males became 0.5 and that of carrier females was 0.8, then from equations (1) and (3) the equilibrium frequencies of affected males and carrier females would be 3.5 μ and 6.3 μ, respectively.

The formulas can be extended to deal with different mutation rates in males (μ_m) and females (μ_f). The contribution of mutation to M is μ_f and to F is $(\mu_m + \mu_f)$. The solution for F then becomes $[\mu_m + \mu_f(1 + m)]/[1 - (f/2)(1 + m)]$, and similarly for M.

COMPLEX EQUILIBRIA

In practice, many other factors are likely to affect the net reproductive fitness of affected males and carrier females, so that the equilibrium frequency will usually be a complex function of several factors. Each of these will first be considered separately. Then the joint effect of the different factors will be considered, and general expressions from which the complex equilibria can be calculated will be derived.

Stage of Detection

If there is a previous family history of a disease or if carrier tests are available, carriers may be detected before any affected offspring are born. Following Fraser [2], this will be called prospective detection. However, most carriers of X-linked conditions will be detected only after the birth (or diagnosis) of an affected son; that is, detection is retrospective. Note that a proportion of carrier females will remain undetected since they have no affected offspring.

Among all carriers not detected prospectively, the proportion of cases born (or preventable) after the first case can be derived as follows. With family size n, an average of $n/4$ affected sons is expected. In a proportion $(3/4)^n$ of families of carrier females, there will be no affected sons. In the remainder $[1 - (3/4)^n]$, there will be at least one affected son, the first case in the family. Thus, the proportion of first cases among all cases in families of size n from carrier females is

$$z_n = [1 - (3/4)^n]/(n/4), \tag{4}$$

as shown by Fraser [5]. Table 1 illustrates this result for families of size three. The proportion of first cases among all affected is 37/48, which is equal to z_n for $n = 3$. The proportion of cases born after the first case is, of course, $(1 - z_n)$.

Since z_n is dependent on family size (n), the distribution of family size in the population must be taken into account. Fraser [5] has evaluated the weighted mean value ($\bar{z}$) for different distributions of family size. For example, for a Poisson

TABLE 1

PROPORTION OF FIRST CASES IN FAMILIES OF SIZE THREE

Family Order 1-2-3	Frequency (× 64)	No. First Cases	No. Affected	No. Normal Born before First Case	Total Normal Born
*NNN**	27	0	0	3	3
NNA	9	1	1	2	2
NAN	9	1	1	1	2
ANN	9	1	1	0	2
NAA	3	1	2	1	1
ANA	3	1	2	0	1
AAN	3	1	2	0	1
AAA	1	1	3	0	0
Total	64	37	48	111	144

* N = normal; A = affected.

distribution with mean family sizes of two and three, the proportions of first cases among all cases are 79% and 70%, respectively. With the negative binomial, the other distribution commonly used to describe distribution of family size, the figures are somewhat higher [5].

It can be shown that $\bar{z}$ also measures the proportion of normal individuals born before the first case. For example, in table 1 this proportion is 111/144. The quantity $\bar{z}$ also gives the average fitness of carrier females if they have no further children after the first case in their family.

Reproductive Fitness

The relative fitness of carrier females will depend on their reproductive practices after detection. Some may terminate their family, others may have normal family size, and still others may compensate for the birth of affected children. A proportion of all carrier females in the population will not be detected, and these are assumed to have normal family size. This group is included implicitly in all the results derived below.

Carrier females who are detected prospectively may have no children and will have a zero fitness. Any who partially restrict their family after detection will have a fitness of less than one, while those who go on to have their normal family size will have a fitness of one. Those terminating their families after the birth of an affected child will have a mean fitness of z.

Some carrier females may compensate for the birth of a affected children so as to have the intended number n of normal children, so-called full reproductive compensation. The total number of children born will be $(a + n)$. The ratio $a/(a + n)$ will be equal to p, the segregation ratio, so $n = a(1 - p)/p$. The mean fitness for such carrier females is then $(a + n)/n$, which is equal to $1/(1 - p)$; this result holds for any family size.

Some carrier females may wish for the intended number of living children. If a

proportion d of affected children do not survive, the total number of children born will be $(n + ad)$. By repeating the above argument, the fitness of such carrier females can be shown to be $1/(1 - dp)$.

Selective Abortion

To deal with selective abortion, the class of offspring selectively aborted must be considered. Various classes of offspring could be selectively aborted: (1) affected males, (2) all males, or (3) affected males and carrier females. Consider the reproductive fitness f^* of carrier females detected prospectively, measured in terms of number of offspring conceived (omitting natural abortions). The observed fitness f, in terms of numbers born, is then

$$f = f^*(x_{NF} + x_{NM} + x_{CF} + x_{AM})/4, \tag{5}$$

where x_{NF}, x_{NM}, x_{CF}, and x_{AM} refer, respectively, to the proportions born of normal females, normal males, carrier females, and affected males conceived. Thus the value of f^* can be derived. Similarly for affected males, female offspring (all carriers) could be selectively aborted. The observed fitness m in terms of offspring born is then

$$m = m^*(2y_{NM} + 2y_{CF})/4, \tag{6}$$

where y_{NM} and y_{CF} are the proportions born of normal males and carrier females conceived.

If the family is detected retrospectively, a proportion z of offspring will be born before detection, as shown in the previous section. The observed fitness f of carrier females in terms of offspring born then becomes

$$f = z + (f^* - z)(x_{NF} + x_{NM} + x_{CF} + x_{AM})/4. \tag{7}$$

Combined Equilibria

All the factors considered above can now be combined into a single set of formulas covering a wide range of situations and can deal with the various factors either singly or in combination.

As discussed earlier, all carrier females are unlikely to adopt the same reproductive practices after detection. To deal with this, let the subscript i refer to the ith group of carrier females who make up a proportion P_i of all carrier females. Similarly, let the subscript j and P_j refer to a particular group of affected males. The procedure then is to calculate, for each group i, the fitness f^*_i in terms of conceptions (given the observed fitness f_i in terms of births) and to weight the values according to the proportion in the group. The equilibrium frequencies can then be written as for equations (1) and (2), namely,

$$M = \mu + \frac{F}{2} \sum_i P_i[z_i + (f^*_i - z_i)x_{AM_i}]; \tag{8}$$

$$F = 2\mu + \frac{F}{2} \sum_i P_i[z_i + (f^*_i - z_i)x_{CF_i}] + M \sum_j P_j m_j y_{CF_j}. \tag{9}$$

By substituting in these formulas, the equilibrium frequencies can be easily derived. If there is no selective abortion, then $f^*_i = f_i$, and all x and $y = 1$. The formulas can be extended to deal with different mutation rates in males and females as before.

Application

To illustrate an application of the previous sections, consider a rather complex case—deriving the equilibrium frequency for an X-linked condition, say hemophilia. Affected males now tend to survive longer and may have more offspring than previously. Genetic counseling is available, and carrier tests and accurate diagnosis can be made. Selective abortion of males from carrier females and of daughters of affected males is also possible and may be used in a proportion of families.

Details of an example are given in tables 2 and 3. Among affected males, suppose 10% do not survive to reproduction, a further 20% have no offspring, and 60% have normal fitness. Suppose the final 10% of the affected males opt for selective abortion of their female offspring, half with no reproductive compensation and half with full reproductive compensation. However, since affected males do not pass X-linked genes to their sons and, using selective abortion, have no daughters, their actual fitness in this case need not be considered.

Among carrier females, those detected prospectively and those detected retrospectively must be treated separately (table 3). For prospective detection the details are similar to those for affected males. If carrier females terminate their family retrospectively, their average fitness ($\bar{z}$) will be about 70%–80%, as discussed earlier. If they use selective abortion of males but have the normal number of pregnancies, the relative fitness for conceptions (f^*) will be 1.0, and for offspring born (f) it will be 0.88. With selective abortion of males and normal family size, the fitness for numbers born (f) will be 1.0. The fitness for numbers conceived (f^*) is then given by $f = 1.0 = 0.75 + (f^* - 0.75)(1 + 1 + 0 + 0)/4$, so that f^* is 1.25. With selective abortion of males and a full family (n) of unaffected

TABLE 2

POSSIBLE REPRODUCTIVE PRACTICES FOR AFFECTED MALES SHOWING PROPORTION AND FITNESS FOR EACH TYPE

AFFECTED MALES	PROPORTION (P_j)	RELATIVE FITNESS	
		No. Born (m_j)	No. Conceived (m^*_j)
Not surviving	0.10	0	0
No offspring	0.20	0	0
Normal family size	0.60	1.0	1.0
Selective abortion of female offspring (n offspring conceived)	0.05	0.5	1.0
Selective abortion of female offspring (n offspring born)	0.05	1.0	2.0

TABLE 3

POSSIBLE REPRODUCTIVE PRACTICES FOR CARRIER FEMALES SHOWING PROPORTION AND FITNESS FOR EACH TYPE

A. DETECTION

	DETECTION OF CARRIER FEMALES	
	Prospective	Retrospective
Proportion	0.20	0.80
Proportion of offspring born before detection ($\bar{z}$)	0.0	0.75

B. REPRODUCTIVE PRACTICE

REPRODUCTIVE PRACTICE AFTER DETECTION	PROSPECTIVE PROPORTIONS (P_i)	FITNESS No. Born (f_i)	FITNESS No. Conceived (f^*_i)	RETROSPECTIVE PROPORTIONS (P_i)	FITNESS No. Born (f_i)	FITNESS No. Conceived (f^*_i)
No further offspring	0.30	0	0	0.40	0.75	0.75
Normal family size (n offspring born)	0.10	1.0	1.0	0.10	1.0	1.0
Selective abortion of males (n offspring conceived)	0.40	0.5	1.0	0.25	0.88	1.0
Selective abortion of males (n offspring born)	0.10	1.0	2.0	0.10	1.0	1.25
Selective abortion of males (n unaffected offspring born)	0.10	1.0	2.0	0.15	1.19	1.63

children, the relative fitness for offspring born (f) is equal to $1 + z/4$, since one-fourth of the children at detection will be affected, and the relative fitness for offspring conceived (f^*) is 1.63.

The values in tables 2 and 3 can be substituted directly in equations (8) and (9) to derive the equilibrium frequencies. These were found to be 3.7 μ for males and 8.4 μ for females. A computer program was written to derive the equilibrium frequencies for any set of variables studied and is available on request.

Relative Importance of Factors

One reason for trying to combine various factors affecting an equilibrium was to study their relative effects and to see where opposing forces would balance. A large variety of comparisons could be made using the above procedures, but only a few cases of special interest are considered here.

In the absence of selective abortion, the critical factor in determining the equilibria is the average fitness of carrier females, and similarly the average fitness of affected males. Thus, the combined effects of any group of fitnesses can be simply

summarized by taking the average fitness for the group and substituting in equation (3) to get the equilibrium frequency.

With selective abortion, the type of offspring born, as well as their number, must be considered. So the average fitness does not summarize the effects of the various factors, and the full procedure as summarized by equations (7), (8), and (9) must be used. The effects of selective abortion with full reproductive compensation are of special interest, for these could theoretically give rise to a continuous increase in frequency with time. However, this is unlikely in practice, for there are other factors with balancing effects. The effects of a proportion of carriers having no children after detection are considered in figure 1. As this proportion

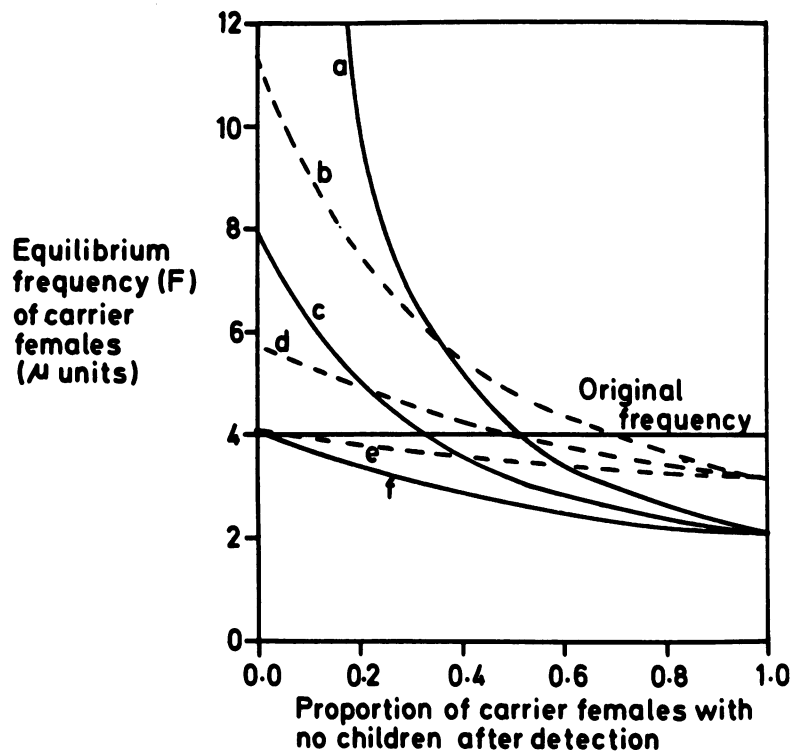

FIG. 1.—Equilibria from selective abortion of all males from carrier females following detection. (Affected males do not reproduce.) *Solid line* = prospective detection; *broken line* = retrospective detection; ($z = 0.7$).

increases the equilibrium frequency in each case falls. Case *a* shows the extreme situation with prospective detection and full reproductive compensation. If only a small proportion of detected carriers have no children, the frequency will not rise continuously but will reach a new, though elevated, equilibrium value.

With retrospective detection and full reproductive compensation (case *b*), much lower equilibrium frequencies are expected. However, a reduction in the number

of births from detected carriers will be expected in practice, since with selective abortion of males, two pregnancies are required on average for each female born. If carrier females have only their intended number of pregnancies, the equilibrium frequencies will be unchanged, or decreased if some carriers have no children after detection (cases *e* and *f*). The two other cases (fig. 1, *c* and *d*) represent intermediate situations. Case *d* refers to retrospective prevention where the intended number of offspring, including the affected proband, is produced. Case *c* represents prospective prevention where half the carriers practice full reproductive compensation and the other half have only their intended number of pregnancies. These result in only moderate increases in frequency above the original equilibrium value.

DISCUSSION

The equilibrium frequencies for X-linked diseases depend on many factors associated with survival and with reproductive practice of affected and carrier individuals. Several workers, but especially Fraser [2], have discussed the effects of changes in single factors acting in isolation, that is, assuming that other factors remain unchanged. The estimated changes in frequency have been very relevant in setting upper limits for the effect of each factor. In practice, several factors are likely to change concurrently, so that to estimate new equilibrium levels more accurately, methods to take the various factors into account are required. The methods outlined here allow combination of any set of factors in order to study and assess their combined equilibrium.

Though the theoretical derivation of equilibria is clear, their determination in practice may be more difficult because the statistics required may be difficult to estimate reliably and will vary over time and place. If the disease is rare, the numbers in various reproductive groups may be small and variable and thus contribute to unreliable estimates. The distributions of intended and completed family size will only be known in retrospect, after reproduction has ceased and may also vary with time. Similarly, the fitness of an individual cannot be measured until the end of reproductive life. These derivations also assume that the disease is caused by a single genetic entity. Both genetic heterogeneity and the existence of phenocopies are factors that would affect the equilibrium values for the frequency of the disease.

With X-linked conditions the approach to new equilibrium levels is reasonably rapid [6], reaching halfway to the new equilibrium in 3–6 generations, compared with hundreds of generations for autosomal recessive conditions. However, it is likely that survival, reproduction, detection, selective abortion, and other factors will change with time and will differ between countries, social groups, and so on. Thus, the equilibria are unlikely to be constant but will continue to vary over time and place.

SUMMARY

Formulas are developed to calculate the equilibrium frequency of X-linked diseases in complex situations. These include the combined effects of survival of

affected individuals, variable reproductive performance of affected males and carrier females, variable stages of detection, and selective abortion. The net effect of any combined set of factors on the equilibrium frequency of the disease can thus be studied.

APPENDIX

Symbols

NF, NM, CF, AM = normal females, normal males, carrier females, affected males
F = equilibrium frequency of CF born
M = equilibrium frequency of AM born
p = segregation ratio
μ = mutation rate per gamete per generation
n = family size
a = average number of AM
z = proportion of offspring born before detection of the family

Parameters Varying with Reproductive Class

P_i = proportion of CF in class (i)
P_j = proportion of AM in class (j)
f_i = relative fitness of CF in class (i) in terms of offspring born
m_j = relative fitness of AM in class (j) in terms of offspring born
f^*_i = relative fitness of CF in class (i) in terms of offspring conceived
m^*_j = relative fitness of AM in class (j) in terms of offspring conceived
x_{NF_i} = proportion of NF born among those conceived by CF in class (i) after detection
y_{NM_j} = proportion of NM born among those conceived of AM in class (j)

REFERENCES

1. HALDANE JBS: The rate of spontaneous mutation of a human gene. *J Genet* 25:251–255, 1935
2. FRASER GR: Selective abortion, gametic selection, and the X chromosome. *Amer J Hum Genet* 24:359–370, 1972
3. MOTULSKY AG, FRASER GR, FELSENSTEIN J: Public health and long-term implications of intrauterine diagnosis and selective abortion. *Birth Defects: Orig Art Ser* 7(5):22–32, 1971
4. EMERY AEH, NELSON MM, MAYO O: Antenatal diagnosis and the muscular dystrophies, in *Actualités de pathologie neuro-musculaire,* edited by SERRATRICE G, Paris, Expansion Scientifique, 1971, pp 13–18
5. FRASER GR: The short-term reduction in birth incidence of recessive diseases as a result of genetic counselling after the birth of an affected child. *Hum Hered* 22:1–6, 1972
6. MORTON NE: Population genetics and disease control. *Social Biol* 18:243–251, 1971